

Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data

David d'Enterria^{1,2} and Juan Rojo³

¹ *CERN, PH Department, CH-1211 Geneva 23, Switzerland*

² *ICREA & ICC-UB, Universitat de Barcelona, 08028 Barcelona, Catalonia*

³ *CERN, PH Department, TH Unit, CH-1211 Geneva 23, Switzerland*

Abstract

The impact of isolated-photon data from proton-(anti)proton collisions at RHIC, Sp̄pS, Tevatron and LHC energies, on the parton distribution functions of the proton is studied using a recently developed Bayesian reweighting method. The impact on the gluon density of the 35 existing isolated- γ measurements is quantified using next-to-leading order (NLO) perturbative QCD calculations complemented with the NNPDF2.1 parton densities. The NLO predictions are found to describe well most of the datasets from 200 GeV up to 7 TeV centre-of-mass energies. The isolated-photon spectra recently measured at the LHC are precise enough to constrain the gluon distribution and lead to a moderate reduction (up to 20%) of its uncertainties around fractional momenta $x \approx 0.02$. As a particular case, we show that the improved gluon density reduces the PDF uncertainty for the Higgs boson production cross section in the gluon-fusion channel by more than 20% at the LHC. We conclude that present and future isolated-photon measurements constitute an interesting addition to coming global PDF analyses.

Contents

1	Introduction	2
2	Experimental data	4
3	Theoretical setup	7
3.1	Isolated-photon cross sections	7
3.2	PDF reweighting	9
4	Results	10
4.1	Comparison between data and NLO pQCD	10
4.2	Sensitivity of isolated- γ to the proton PDFs	18
4.3	Impact of isolated-photon data on the gluon	19
5	Predictions of the gluon-fusion Higgs boson production cross section	24
6	Summary and outlook	26

1 Introduction

The accurate determination of the parton distribution functions (PDFs) of the proton in a wide range of momentum fractions x and energy scales Q [1] is a crucial ingredient for precision phenomenology at the Large Hadron Collider (LHC) [2,3]. Among all parton distributions, the gluon density $g(x, Q^2)$ is one of the least constrained PDFs since it does not couple directly to the photon in deep-inelastic scattering (DIS) measurements of the proton F_2 structure function [4]. A precise determination of $g(x, Q^2)$ is of paramount importance for the LHC physics program both in standard model (SM) as well as in new physics searches. Indeed, gluons drive a significant fraction of the scattering processes at the LHC and gluon-gluon fusion is the dominant channel for the production of the SM Higgs boson [5,6], top-quark pairs or dijets, to mention a few.

In global PDF analyses, the gluon density is directly constrained mostly by jet production and indirectly constrained by scaling violations of $F_2(x, Q^2)$ in DIS and by the momentum sum rule. It is thus important to find new collider observables that provide independent information on $g(x, Q^2)$ and help improve the accuracy of its determination and reduce its associated uncertainties. Prompt photon production in hadronic collisions, defined as the production of photons not issuing from the electromagnetic decays of hadrons, appears as an excellent observable to determine the gluon PDF, since at leading order (LO) it probes the gluon directly through the quark-gluon “Compton” process $qg \rightarrow \gamma q$ [7–9]. As a matter of fact, up to about 12 years ago, prompt photon data were used to constrain $g(x, Q^2)$ in global PDF fits [10,11], as well as in independent determinations of the strong coupling constant [12]. However, a series of measurements carried out at centre-of-mass energies $\sqrt{s} \approx 20\text{--}40$ GeV by the fixed-target E706 experiment [13–15] showed an enhanced γ cross section compared to next-to-leading-order (NLO) perturbative quantum chromodynamics (pQCD) predictions [16–18]. The theoretical cross section deficit was only partially cured by the inclusion of extra soft-gluon resummation contributions [19–23]. Such data-theory discrepancies at fixed-target energies, together with the availability of more precise jet measurements from the Tevatron, lead to abandon the use of inclusive photon data in PDF analyses.

The last PDF parametrisation that used the prompt photon data was MRST99 [11]. Since then, the issue of the compatibility of NLO pQCD with photon measurements has been discussed extensively. For example, the authors of Ref. [24] performed a systematic comparison of NLO calculations with most of the available photon results by 2006 (see also Refs. [25, 26] for previous studies) and showed that collider *isolated* photon data was in reasonable agreement with the theoretical predictions. Indeed, as discussed in detail in [27], by (i) increasing the \sqrt{s} from fixed-target to collider energies, and by (ii) requiring the photon to be isolated from any hadronic activity within a given distance around its direction, one is left with an isolated-photon sample dominated by energy scales away from non-perturbative effects (such as intrinsic- k_T broadening [28]), and where the uncertain contribution from photons issuing from the collinear fragmentation of final-state partons [29] are significantly reduced. Figure 1 shows the relative contribution of the three leading diagrams – direct Compton, direct quark-antiquark annihilation, and fragmentation – for inclusive (left) and for isolated (right) photon production in proton-antiproton ($p\bar{p}$) at Tevatron (top) and in proton-proton ($p\text{-}p$) collisions at the nominal LHC energy (bottom). The fragmentation- γ component, including the poorly known gluon-to-photon fragmentation function [30, 31] which dominates the low E_T^γ part of the inclusive photon spectra, is significantly reduced after application of isolation cuts and accounts only for less than 10–15% of the cross sections at both energies. The right panels of Fig. 1 clearly indicate that the Compton component is the dominant contribution to the isolated- γ spectra in wide ranges of the measured transverse energy E_T^γ , in particular for $p\text{-}p$ collisions, thus guaranteeing their increased sensitivity to $g(x, Q^2)$.

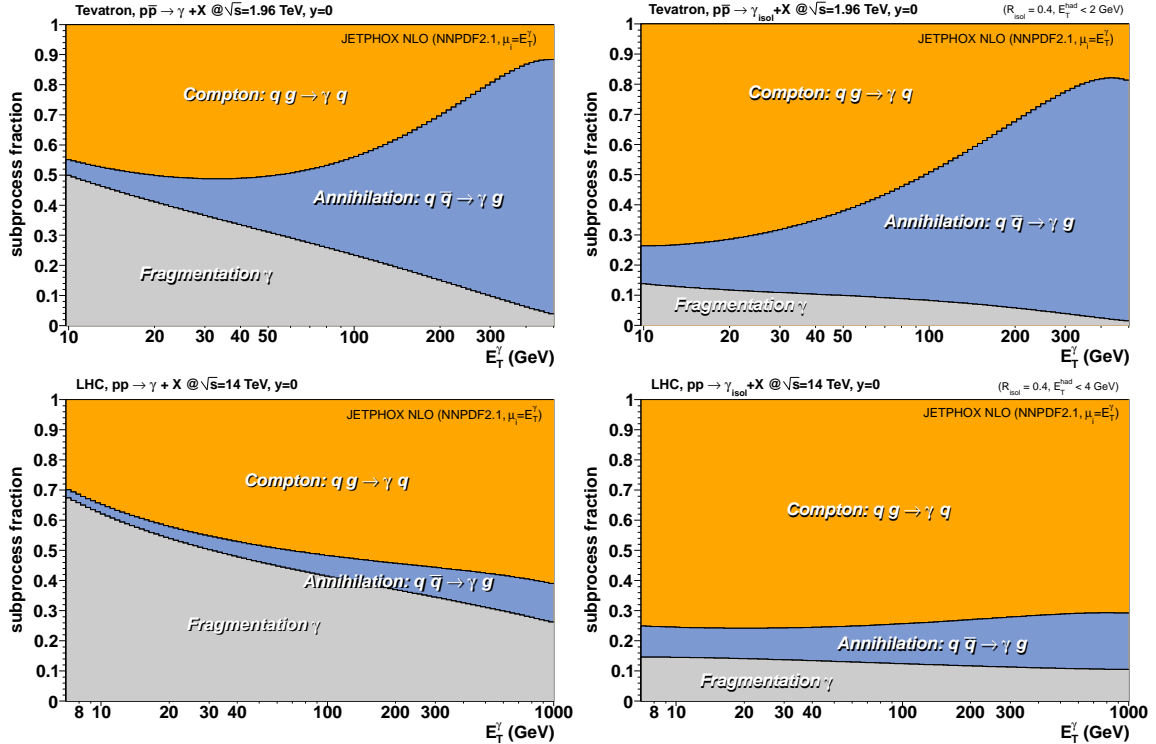


Figure 1: Relative contributions of the qg -Compton, $q\bar{q}$ -annihilation and fragmentation subprocesses in inclusive (left) and isolated (right) prompt-photon production in $p\bar{p}$ (Tevatron at 1.96 TeV, top) and $p\text{-}p$ (LHC at 14 TeV, bottom) collisions at midrapidity. Results obtained at NLO with JETPHOX [32] using NNPDF2.1 [33], theoretical scales set to $\mu = E_T^\gamma$, and BFG-II FFs [31] (see Sect. 3).

In view of the various accurate measurements of isolated-photon spectra carried out recently at the LHC [34–38], which have added up about 130 new points to the existing world data, it is a timely moment to revisit first if NLO pQCD provides a good description of all available isolated- γ collider data, and if so, what quantitative constraints these new results impose on the gluon PDF. However, the calculation of cross sections at NLO accuracy in hadronic collisions is usually a very time-consuming process as it involves the Monte Carlo (MC) generation of a large number of event weights for the integration over the hard subprocess phase-space in order to cancel infrared divergences. This fact makes the full-NLO calculation with different PDF choices, as in iterative fits of the parton densities, usually prohibitive in terms of the required processing time. Computational alternatives exist, though, such as e.g. (i) the APPLGRID framework [39] or the FASTNLO software [40], both of which make use of lookup tables with cross-section weights which can be *a posteriori* fast-convoluted with any parton density set, and (ii) the NNPDF Bayesian reweighting technique [41, 42] which can be applied to any set of MC-based parton densities and which relies on the computation of the χ^2 distribution between the data and each MC replica in order to determine the new (reweighted) PDFs.

In this paper we use the NNPDF reweighting method in order to quantify the impact of collider isolated- γ data on $g(x, Q^2)$. Theoretical NLO pQCD predictions obtained with the JETPHOX program [32] are compared to all available experimental data in a very wide range of centre-of-mass (c.m.) energies ranging from 200 GeV up to 7 TeV. These comparisons are then used to quantitatively determine to which extent isolated-photons can be used to reduce the gluon PDF uncertainties.

The paper is organised as follows. In Sect. 2 the experimental isolated-photon results used in the analysis, more than 30 measurements with a total of $\mathcal{O}(400)$ data points, are presented. Section 3 describes the setup of the calculation based on the the JETPHOX code to provide NLO pQCD predictions and on the NNPDF-reweighting method of the parton densities. The main results of this analysis, the data–theory comparisons and the quantitative impact of the photon data on the reweighted $g(x, Q^2)$, are discussed in Sect. 4. In Sect. 5, the predictions for Higgs cross sections in the gluon-fusion production channel using the improved gluon distributions are presented. We conclude in Sect. 6 with a summary of the main results and an outline of possible future developments regarding the use of isolated-photon data in global PDF fits.

2 Experimental data

The world isolated- γ data covers all prompt-photon measurements at collider machines in the last 30 years. Indeed, for c.m. energies above a hundred of GeV, photon isolation is experimentally required in order to identify the high- E_T prompt signal out of the overwhelming background of photons from the decays of π^0 and η mesons produced in the fragmentation of jets. Table 1 summarises the details of the 35 existing experimental measurements. The table lists the characteristics of the more than 400 isolated- γ data-points measured in p - p and p - \bar{p} collisions from RHIC c.m. energies ($\sqrt{s} = 200$ GeV) up to the highest energies available so far at the LHC ($\sqrt{s} = 7$ TeV). The table updates the results collected in [43] by including the latest ATLAS [35] and CMS [38] measurements. For each system we provide the kinematical coverage in E_T^γ and rapidity y_γ of the measured photon, the corresponding range of parton fractional momentum x covered by the measurement, as well as the applied photon isolation criteria used in each case (maximum hadronic energy E_h , or fraction of the photon energy ε_h , allowed within the isolation

cone of radius $R = \sqrt{(y_h - y_\gamma)^2 + (\phi_h - \phi_\gamma)^2}$ around the photon direction).

System	Collab./Experiment (collider) [Ref.]	\sqrt{s} (TeV)	$ y_\gamma $ range	E_T^γ range (GeV)	x range	Data points	Isolation radius, had. energy
p - p	ATLAS (LHC) [34]	7.	<0.6	15–100	$5 \cdot 10^{-3}$ –0.05	8	$R = 0.4, E_h < 5$ GeV
p - p	ATLAS (LHC) [34]	7.	0.6–1.37	15–100	$3 \cdot 10^{-3}$ –0.1	8	$R = 0.4, E_h < 5$ GeV
p - p	ATLAS (LHC) [34]	7.	1.52–1.81	15–100	$2 \cdot 10^{-3}$ –0.1	8	$R = 0.4, E_h < 5$ GeV
p - p	ATLAS (LHC) [35]	7.	<0.6	45–400	$5 \cdot 10^{-3}$ –0.1	8	$R = 0.4, E_h < 4$ GeV
p - p	ATLAS (LHC) [35]	7.	0.6–1.37	45–400	$5 \cdot 10^{-3}$ –0.2	8	$R = 0.4, E_h < 4$ GeV
p - p	ATLAS (LHC) [35]	7.	1.52–1.81	45–400	$2 \cdot 10^{-3}$ –0.3	8	$R = 0.4, E_h < 4$ GeV
p - p	ATLAS (LHC) [35]	7.	1.81–2.37	45–400	$2 \cdot 10^{-3}$ –0.5	8	$R = 0.4, E_h < 4$ GeV
p - p	CMS (LHC) [37]	7.	<1.45	21–300	$5 \cdot 10^{-3}$ –0.1	11	$R = 0.4, E_h < 5$ GeV
p - p	CMS (LHC) [36]	7.	<0.9	25–400	$5 \cdot 10^{-3}$ –0.2	15	$R = 0.4, E_h < 5$ GeV
p - p	CMS (LHC) [36]	7.	0.9–1.44	25–400	$2 \cdot 10^{-3}$ –0.3	15	$R = 0.4, E_h < 5$ GeV
p - p	CMS (LHC) [36]	7.	1.57–2.1	25–400	10^{-3} –0.4	15	$R = 0.4, E_h < 5$ GeV
p - p	CMS (LHC) [36]	7.	2.1–2.5	25–400	10^{-3} –0.5	15	$R = 0.4, E_h < 5$ GeV
p - p	CMS (LHC) [38]	2.76	<1.45	20–80	10^{-3} –0.05	6	$R = 0.4, E_h < 5$ GeV
p - \bar{p}	CDF (Tevatron) [44]	1.96	<1.0	30–400	0.01–0.4	16	$R = 0.4, \varepsilon_h < 0.1$
p - \bar{p}	D0 (Tevatron) [45]	1.96	<0.9	23–300	0.01–0.3	17	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	CDF (Tevatron) [46]	1.8	<0.9	11–132	$5 \cdot 10^{-3}$ –0.2	17	$R = 0.4, E_h < 4$ GeV
p - \bar{p}	CDF (Tevatron) [47]	1.8	<0.9	10–65	$5 \cdot 10^{-3}$ –0.1	17	$R = 0.4, E_h < 1$ GeV
p - \bar{p}	CDF (Tevatron) [48]	1.8	<0.9	8–132	$5 \cdot 10^{-3}$ –0.2	16	$R = 0.7, E_h < 2$ GeV
p - \bar{p}	D0 (Tevatron) [49]	1.8	<0.9	10–140	$5 \cdot 10^{-3}$ –0.2	9	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	D0 (Tevatron) [49]	1.8	1.6–2.5	10–140	10^{-3} –0.4	9	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	D0 (Tevatron) [50]	1.8	<0.9	9–126	$5 \cdot 10^{-3}$ –0.2	23	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	D0 (Tevatron) [50]	1.8	1.6–2.5	9–126	10^{-3} –0.4	23	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	CDF (Tevatron) [46]	0.63	<0.9	8–38	0.01–0.2	7	$R = 0.4, E_h < 4$ GeV
p - \bar{p}	D0 (Tevatron) [51]	0.63	<0.9	7–50	0.01–0.3	7	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	D0 (Tevatron) [51]	0.63	1.6–2.5	7–50	10^{-3} –0.4	7	$R = 0.4, E_h < 2$ GeV
p - \bar{p}	UA1 (Sp̄pS) [52]	0.63	<0.8	16–100	0.03–0.3	16	$R = 0.7, E_h < 2$ GeV
p - \bar{p}	UA1 (Sp̄pS) [52]	0.63	0.8–1.4	16–70	0.01–0.4	10	$R = 0.7, E_h < 2$ GeV
p - \bar{p}	UA1 (Sp̄pS) [52]	0.63	1.6–3.0	16–70	0.01–0.5	13	$R = 0.7, E_h < 2$ GeV
p - \bar{p}	UA2 (Sp̄pS) [53]	0.63	<0.76	14–92	0.03–0.3	13	$R = 0.265, \varepsilon_h < 0.25$
p - \bar{p}	UA2 (Sp̄pS) [54]	0.63	<0.76	12–83	0.03–0.3	14	$R = 0.25, E_h < 0.1$ GeV
p - \bar{p}	UA2 (Sp̄pS) [54]	0.63	1.0–1.8	12–51	0.01–0.4	8	$R = 0.53, E_h < 2$ GeV
p - \bar{p}	UA1 (Sp̄pS) [52]	0.546	<0.8	16–51	0.03–0.2	6	$R = 0.7, E_h < 2$ GeV
p - \bar{p}	UA1 (Sp̄pS) [52]	0.546	0.8–1.4	16–46	0.02–0.4	5	$R = 0.7, E_h < 2$ GeV
p - \bar{p}	UA1 (Sp̄pS) [52]	0.546	1.6–3.0	16–38	0.01–0.5	5	$R = 0.7, E_h < 2$ GeV
p - p	PHENIX (RHIC) [55]	0.2	<0.35	3–16	0.03–0.2	17	$R = 0.5, \varepsilon_h < 0.1$

Table 1: World systematics of isolated-photon data in p - p and p - \bar{p} collisions. For each system, we quote (i) the experiment, collider and bibliographical reference, (ii) centre-of-mass energy \sqrt{s} , the measured (iii) rapidity $|y_\gamma|$ and (iv) E_T^γ ranges, (v) the parton fractional momenta x probed, (vi) number of data points, and the (v) isolation criteria used.

Since we have decided not to consider the existing prompt- γ measurements below $\sqrt{s} \approx 65$ GeV, both at fixed-target energies and at the CERN-ISR collider (see the compilation [25]), as they deal with *inclusive*- γ cross sections with large dependence on poorly known parton-to-photon FFs and on non-perturbative corrections due to the smaller scales involved, the lowest-energy isolated- γ measurement is that of the PHENIX experiment in p - p at $\sqrt{s} = 200$ GeV at RHIC [55]. Next in ascending order of collision energies are the oldest measurements by the CERN Sp̄pS UA1 [52] and UA2 [53, 54] collaborations at $\sqrt{s} = 546$ and 630 GeV, which amount to a total of 90 data-points. The Tevatron datasets are available for $\sqrt{s} = 0.63$ TeV [46, 51], 1.8 TeV [46–50] and 1.96 TeV [44, 45], the most precise data being the latter. The total number of data-points is 21 at 0.63 TeV, 114 at 1.8 TeV, and 33 at 1.96 TeV. Finally, the recent LHC measurements from ATLAS and CMS amounting to 133 data points [34–38], cover the wider range in rapidity (up to $|y_\gamma| = 2.5$) and have smaller systematical uncertainties ($\pm(7-20)\%$ depending on E_T^γ and y_γ) than all other previous measurements. For ATLAS we consider both the 880 nb $^{-1}$ [34] and 36 pb $^{-1}$ [35]

measurements, while for CMS we include the 7 TeV 2.9 pb⁻¹ [37] and 36 pb⁻¹ analyses [36], as well as the latest p - p measurement carried out at 2.76 TeV [38] for reference heavy-ion collisions studies.

The isolated-photon data of Table 1 are plotted in Fig. 2 as a function of E_T^γ (left panel) and as function of $x_T = 2E_T^\gamma/\sqrt{s}$ with the cross sections scaled by \sqrt{s}^n (right panel). All spectra in Fig. 2 (left) follow clear power-law dependencies from 3 to 400 GeV spanning 9 orders of magnitude in the cross sections. The x_T spectra (Fig. 2, right) coalesce over a single curve when the cross sections are normalised by \sqrt{s}^n with exponent $n \approx 4.5$. Such a behaviour is very close to the $1/p_T^{n=4}$ dependence expected for partonic $2 \rightarrow 2$ scattering cross sections in the conformal QCD limit, disregarding scaling violations from PDF and running of α_s [56]. A few deviations are visible, in particular for measurements that are either away from midrapidity for which the assumption $x_T = 2E_T^\gamma/\sqrt{s}$ does not exactly hold and/or for which the highest photon energies are mis-reconstructed (see Sect. 4.1). The fact that all data satisfy the same scaling with \sqrt{s} and collapse into a single curve indicates the universality of the underlying partonic production mechanism. The observed E_T^γ power-laws and universal \sqrt{s} -dependence of the cross sections are tell-tale of a perturbative origin for the production of isolated photons and confirm the validity of pQCD calculations to study their yields.

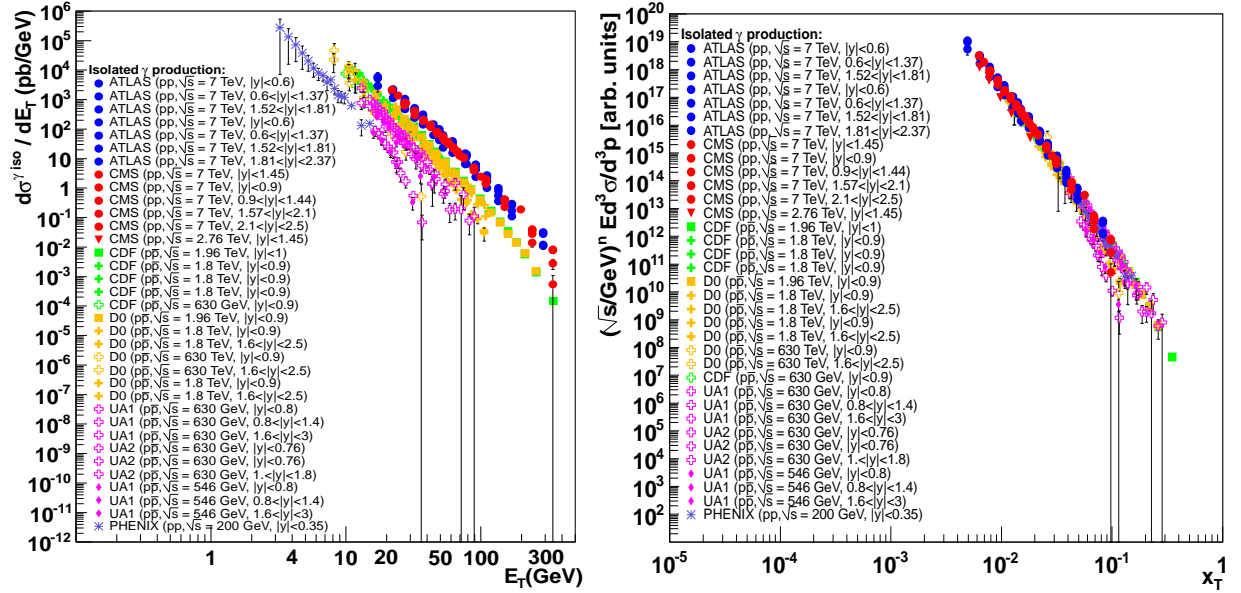


Figure 2: World systematics of isolated-photon spectra measured in p - p and p - \bar{p} collisions at collider energies (Table 1) as a function of E_T^γ (left) and x_T (right) where the invariant cross sections have been scaled by \sqrt{s}^n with $n = 4.5$.

In Fig. 3 we show a scatter plot of the world isolated-photon measurements listed in Table 1 together with the DIS, Drell-Yan and jet data (about 3500 points) used in the NNPDF2.1 global fits. For each experimental data-point we plot two values of parton fractional momentum $x_\pm = x_T \cdot e^{\pm y_\gamma}/\sqrt{s}$, assuming leading-order partonic kinematics. The plot shows that the photon LHC data extend the kinematic coverage in particular over the region $x \approx 10^{-3} - 10^{-2}$ at moderately large energy scales ($E_T^{\gamma^2} \approx 10^3 - 10^5$ GeV²) not directly covered by the other experimental datasets.

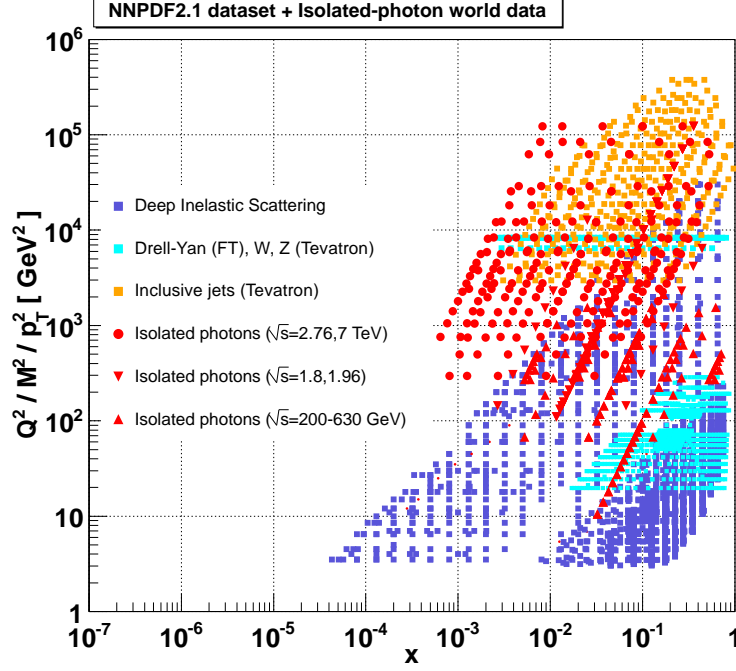


Figure 3: Kinematical region in the $x - Q^2$ plane probed by experimental isolated- γ data at collider energies (red circles and triangles) which enter into this analysis (Table 1) compared to the coverage of DIS, Drell-Yan and jet datasets (squares) used in the NNPFD2.1 global fits.

3 Theoretical setup

In this section the basic ingredients of the JETPHOX program used to compute the isolated-photon cross sections are discussed, and the NNPFD reweighting technique employed to quantify the impact of new data on the proton PDFs is briefly recalled.

3.1 Isolated-photon cross sections

Two types of processes contribute at leading order to prompt photon production in p - p and p - \bar{p} collisions: the ‘direct’ contribution, where the photon is emitted directly from a pointlike coupling to the hard parton-parton vertex, and the ‘fragmentation’ (called also ‘anomalous’ in the past) contribution, in which the photon originates from the collinear fragmentation of a final-state parton. Schematically, the differential photon cross section as a function of transverse energy E_T^γ and rapidity y_γ can be written as

$$d\sigma \equiv d\sigma_{\text{dir}} + d\sigma_{\text{frag}} = \sum_{a,b=q,\bar{q},g} \int dx_a dx_b f_a(x_a; \mu_F^2) f_b(x_b; \mu_F^2) \times \quad (1)$$

$$\left[d\hat{\sigma}_{ab}^\gamma(p_\gamma, x_a, x_b; \mu_R, \mu_F, \mu_{\text{ff}}) + \sum_{c=q,\bar{q},g} \int_{z_{\min}}^1 \frac{dz}{z^2} d\hat{\sigma}_{ab}^c(p_\gamma, x_a, x_b, z; \mu_R, \mu_F, \mu_{\text{ff}}) D_c^\gamma(z; \mu_{\text{ff}}^2) \right]$$

where $f_a(x_a; \mu_F^2)$ is the parton distribution function of parton species a inside the incoming protons at momentum fraction x_a ; $d\hat{\sigma}_{ab}$ are the parton-parton subprocess differential cross sections; and $D_{\gamma/k}(z; \mu_{\text{ff}}^2)$ is the fragmentation function of parton k to a photon carrying a fraction z of the parent parton energy, integrated from $z_{\min} = x_T \cosh y_\gamma$ to 1. The scaled momentum x_T is

a good representative for the typical x values probed in the PDFs at central rapidities, while at forward rapidities one is probing $x \sim x_T \exp(-y_\gamma)$. The arbitrary parameters μ_R , μ_F and μ_{ff} are respectively the renormalisation, initial-state factorisation, and fragmentation scales which encode any residual dependence of the cross sections to higher-order contributions which are missing in the calculation.

The study provided in this article relies on the JETPHOX calculation of both $d\sigma_{\text{dir}}$ and $d\sigma_{\text{frag}}$ at NLO accuracy [32] in the strong coupling $\alpha_s(\mu_R)$, i.e. all diagrams up to the order $\mathcal{O}(\alpha_s^2)$ are included, in the $\overline{\text{MS}}$ renormalisation scheme. The results of the NLO calculation of $d\sigma_{\text{dir}}$ have been known for a long time [57]. The calculation of the NLO corrections to $d\sigma_{\text{frag}}$ became also available later [18, 32, 58]. We note that the distinction between $d\sigma_{\text{dir}}$ and $d\sigma_{\text{frag}}$ is arbitrary and only its sum is physically observable, e.g. bremsstrahlung from a quark leg can be considered as “fragmentation” or as “NLO direct” depending on the value of the fragmentation scale considered.

A few recent works exist that have computed various beyond-NLO corrections to the inclusive production of prompt photons:

- The small- x high-energy corrections [59] were computed in Ref. [60] focusing on the low- E_T^γ region of the spectrum where power terms of the type $\alpha_s^k \ln^p(x)$ are enhanced when the scaling variable x becomes small. It was found that such type of next-to-NLO (NNLO) corrections are negligible in the kinematical range of the available collider data and, thus, ignored in the following.
- Resummation at next-to-leading-logarithmic (NLL) [19–23] or even next-to-next-to-leading (NNLL) [61] accuracy of threshold and recoil contributions due to soft gluon emissions which are large close to the phase space boundary when E_T^γ is about half of the c.m. energy ($x_T \rightarrow 1$) have been obtained for $d\sigma_{\text{dir}}$ and $d\sigma_{\text{frag}}$. The effect of this resummation is important at very large photon E_T corresponding to values $x_T \gtrsim 0.2$, and provide a much reduced scale dependence than the NLO approximation. In the current collider data the typical ranges involved ($x \approx 0.001$ – 0.2 , see Fig. 3) lead to small threshold corrections which we do not consider either.
- At very large photon transverse energies $E_T^\gamma \approx 1$ – 2 TeV, not yet measured at the LHC, one should also consider additional corrections due to electroweak boson exchanges which decrease the photon yields by about 10–20% [62].

In this paper, we use the JETPHOX (version 1.3.0) Monte Carlo program [32] to compute the NLO pQCD predictions for isolated-photon production. We take into account five active quark flavours. The gluon box diagram $gg \rightarrow g\gamma$ is included in the calculations although its contribution to the single inclusive spectrum is found to be just of a few percent. The NNPDF2.1 parton densities (with 100 replicas per system) [33] were interfaced to JETPHOX via the LHAPDF (version 5.8.5) package [63]. The value of $\alpha_s = 0.119$ used in the calculations is that provided by the NNPDF2.1 parametrisation itself. The renormalisation, factorisation and fragmentation scales are all set equal to the photon transverse energy, $\mu_F = \mu_R = \mu_{\text{ff}} = E_T^\gamma$. Such μ scales are found to result in spectra that agree well with the central values of the experimental data. At LHC and Tevatron energies, the sensitivity to changes in the arbitrary theoretical scales was discussed in [27] and found to be of about $\pm(10$ – $15)\%$ in the measured E_T^γ ranges. The parton-to-photon fragmentation functions used are the BFG-II (“large gluon”) set [31]. The isolated- γ spectrum obtained with the alternative BFG-I (“small gluon”) set is just a few percent smaller than the

one obtained with BFG-II in the lowest E_T^γ range [27] since a significant fraction (up to 80%, see Fig. 1) of the fragmentation photons are removed by the application of isolation cuts.

We run JETPHOX for all the systems in Table 1 within the (E_T^γ, y_γ) kinematics ranges of every measurement. The MC photon isolation criteria are also matched as closely as possible to each one of the experimental cuts. We histogram the partonic configurations generated in E_T^γ -bins of the same size as for the experimental data in order to be able to compute the χ^2 bin-by-bin and avoid potential problems with steeply-falling spectra such as those discussed in Ref. [64]. Running for each one of the 100 NNPDF2.1 replicas with acceptable MC statistics (statistical uncertainty below 1%) for all E_T^γ bins is a very CPU-demanding task. The JETPHOX MC production for 100 PDF replicas takes about 7 (resp. 10) hours per each million direct (resp. fragmentation) events in a $\mathcal{O}(2\text{ GHz})$ CPU core, which means that about one week of computer-time is needed to obtain the 100 spectra for each one of the 35 systems considered.

3.2 PDF reweighting

In order to quantify the impact of the collider isolated-photon data on the proton PDFs we use the Bayesian reweighting method described in Refs. [41, 42]. This technique can be applied to any PDF-set that estimates PDF uncertainties based on a MC method, such as the NNPDF family [33, 42, 65–70] (but also MSTW [71] and HERAPDF [72] have produced MC sets). The method allows one to straightforwardly determine the impact of a new dataset on PDFs by means of Bayesian inference. The only ingredient needed are the data–theory “goodness-of-fit” χ_k^2 for each k -th MC replica (where $k = 1, \dots, N_{\text{rep}}$ runs over the full set of replicas: $N_{\text{rep}} = 100$ or 1000 in the NNPDF2.1 set). This technique reduces the problem of the slowness of JETPHOX (or in general any NLO hadronic computation) which forbids its direct use within a PDF-fitting program, since the theoretical predictions need to be computed only once for each PDF replica.

The χ_k^2 for each replica is defined as follows:

$$\chi_k^2 = \frac{1}{N_{\text{dat}}} \sum_{i=1}^{N_{\text{dat}}} \frac{\left(\sigma_i^{(\text{th}), (k)} - \sigma_i^{(\text{exp})} \right)^2}{\Delta_{\text{tot}}^2}, \quad (2)$$

where $\sigma_i^{(\text{th}), (k)}$ is the NLO theoretical prediction for the isolated-photon cross section, Eq. (1), obtained with the f_k PDF replica, $\sigma_i^{(\text{exp})}$ is the corresponding experimental measurement and Δ_{tot} accounts for the experimental uncertainties. N_{dat} is the number of data points of each particular measurement. We note that the experimental covariance matrix is not available for any of the isolated-photon datasets and thus we are forced to add in quadrature the statistical and systematic uncertainties into Δ_{tot} . Likewise, absolute normalisation uncertainties (e.g. from the integrated luminosity) should be in principle included in the experimental covariance matrix using the t_0 method (as consistently done within the NNPDF analysis) [73], but they are typically smaller than other experimental uncertainties and also added in quadrature here.

Once the χ_k^2 for each replica has been computed, the new weight of the replica is given by

$$w_k = \frac{(\chi_k^2)^{\frac{1}{2}(n-1)} e^{-\frac{1}{2}\chi_k^2}}{\frac{1}{N_{\text{rep}}} \sum_{k=1}^N (\chi_k^2)^{\frac{1}{2}(n-1)} e^{-\frac{1}{2}\chi_k^2}}. \quad (3)$$

The weights w_k , when divided by the number of MC replicas of the prior PDF set (N_{rep}), are then simply the probabilities of the replicas f_k , given the χ_k^2 to the new added experimental results. If the new data constrains the PDFs, reweighting will be less efficient than refitting because of the discarded replicas with low weight. One can quantify this efficiency loss by using the Shannon entropy to compute the effective number of replicas left after reweighting:

$$N_{\text{eff}} \equiv \exp \left\{ \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} w_k \ln(N_{\text{rep}}/w_k) \right\}. \quad (4)$$

Clearly $0 < N_{\text{eff}} < N_{\text{rep}}$ and the reweighted fit has the same accuracy as a refit with N_{eff} replicas. This effective number of replicas N_{eff} is a useful estimator of the impact of each individual dataset. The smaller N_{eff} is, the more the new dataset constrains the PDFs.

It is often the situation that experimental uncertainties are under or over estimated. It is then possible to rescale the uncertainties of the data by a factor α , and then use inverse probability to calculate the probability density for the rescaling parameter α

$$\mathcal{P}(\alpha) \propto \frac{1}{\alpha} \sum_{k=1}^{N_{\text{rep}}} w_k(\alpha). \quad (5)$$

Here $w_k(\alpha)$ are the weights Eq. (3) evaluated by replacing χ_k^2 with χ_k^2/α^2 , and are thus proportional to the probability of f_k given the new data with rescaled errors. Averaging $w_k(\alpha)$ in the reweighted fit thus gives the probability density for α . If this probability density peaks close to unity, the new data are consistent with the pre-reweighting data, while if it peaks far above (below) one then it is likely that the errors in the data have been under (over) estimated. The distribution of the rescaling variable α , normalised to unity, is used in the next section to investigate if the bad agreement between NLO pQCD and a few datasets could be due to possibly underestimated experimental uncertainties.

4 Results

The comparison between NLO pQCD and the experimental data for all systems listed in Table 1 is presented in the next subsection. With this information we compute for each case the data-theory χ_k^2 distributions Eq. (2), the weights Eq. (3), the effective number of replicas Eq. (4), the sensitivity to individual PDF flavours, and finally the associated constraints on the gluon distribution.

4.1 Comparison between data and NLO pQCD

Using the JETPHOX program with the setup discussed in Sect. 3.1 we have computed the $E_T^{\gamma\gamma}$ -differential cross sections using the $N_{\text{rep}} = 100$ replicas of the NNPDF2.1 NLO parton distributions for all the systems listed in Table 1. The ratios between data and theory are shown in Figs. 4–9, where the (yellow) band corresponds to the distribution of the predictions obtained which each one of the 100 replicas, while the outer error-bars cover the sum in quadrature of the statistical and systematic uncertainties of the measurement.

Figure 4 shows the data/NLO ratios for the isolated-photons measured at the lowest c.m. energies: $\sqrt{s} = 200$ GeV (RHIC) and 546 GeV (Sp̄pS). The PHENIX results agree well with NLO

pQCD within the quite large experimental uncertainties. For UA1 and UA2, the data/theory ratio is around unity below $E_T^\gamma \approx 25$ GeV, but the highest transverse energy points are clearly overestimated by NLO in two cases, likely due to an experimental mis-reconstruction of the photon energy at the end of the spectra measured at the time (see below).

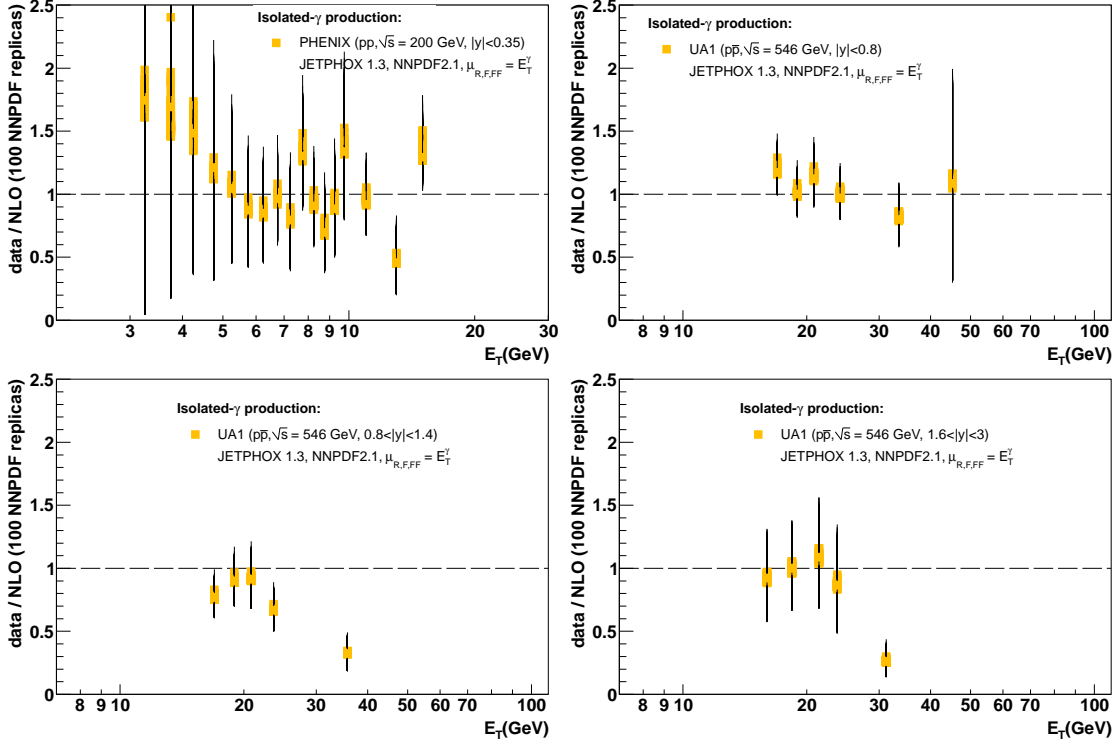


Figure 4: Ratio of isolated-photon data and NLO pQCD predictions for p - p collisions at $\sqrt{s} = 200$ GeV (PHENIX) and p - \bar{p} collisions at $\sqrt{s} = 546$ GeV (UA1) and various photon rapidities. The (yellow) band indicates the range of predictions for each one of the 100 NNPDF2.1 replicas, and the bars show the total experimental uncertainty.

The comparison of the theoretical predictions to the UA1, UA2, CDF and D0 data at 630 GeV c.m. energies and various photon rapidities are shown in Fig. 5. With a few exceptions, a general trend appears indicating that the measured cross-sections are higher (respectively lower) than the NLO calculations at the low (resp. high) end of the photon energy spectrum. The disagreement is larger for increasingly forward rapidities. Yet, given the relatively large experimental uncertainties, theory and data are in general consistent and show χ^2 not very far from one in most of the cases.

The Tevatron Run-I measurements at $\sqrt{s} = 1.8$ TeV are compared to the NLO predictions in Fig. 6. The two oldest measurements [48, 50] (not plotted here) are systematically below not only the pQCD predictions but also the spectra obtained by the same experiments in a basically identical kinematic range a few years later [46, 47, 49]. By discarding those superseded datasets from further analysis, we find a generally good consistency of the 1.8-TeV data with NLO pQCD.

The comparison of NLO to the Tevatron CDF and D0 Run-II results at $\sqrt{s} = 1.96$ TeV is shown in Fig. 7. While at the level of χ^2 the data-theory agreement is good, in both cases

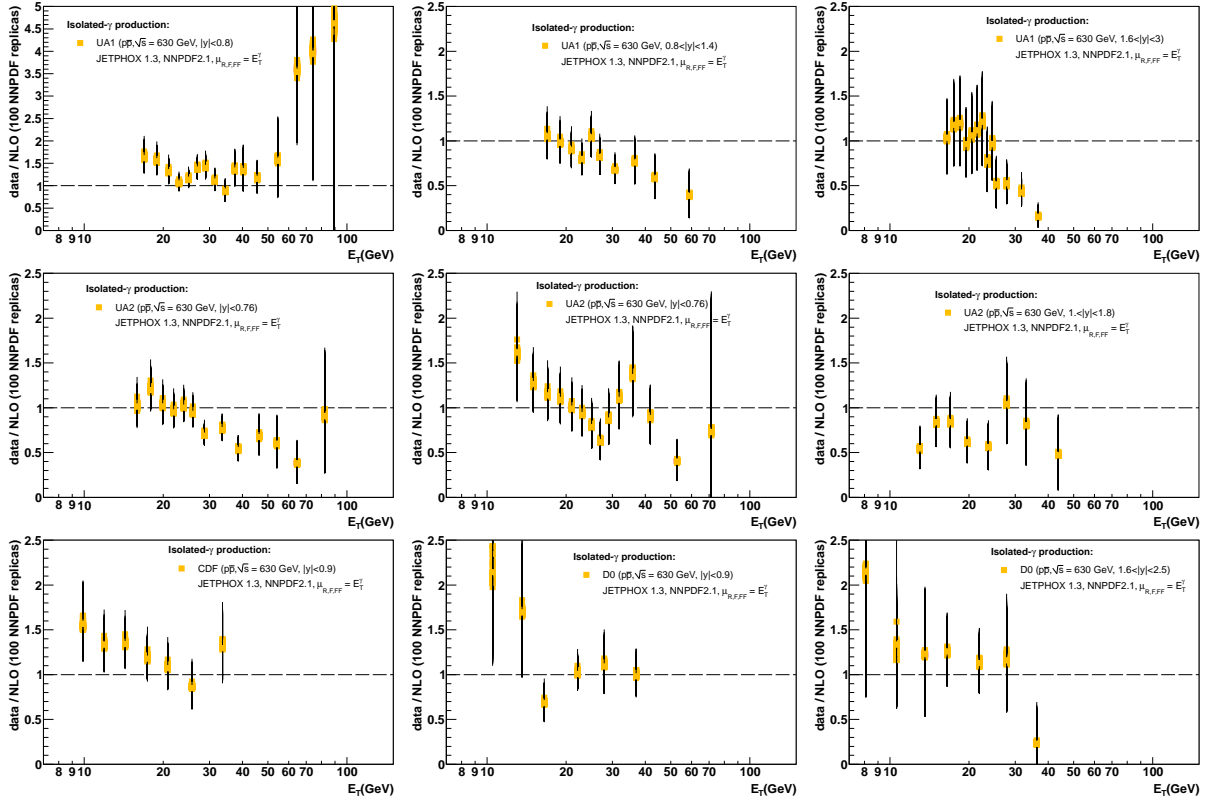


Figure 5: Ratio of isolated-photon data (UA1 and UA2 in the 3 top and middle plots, and CDF and D0 in the 3 bottom panels) and NLO pQCD predictions for $p\bar{p}$ collisions at $\sqrt{s} = 630$ GeV at various photon rapidities. The (yellow) band indicates the range of predictions for each one of the 100 NNPDF2.1 replicas, and the bars show the total experimental uncertainty.

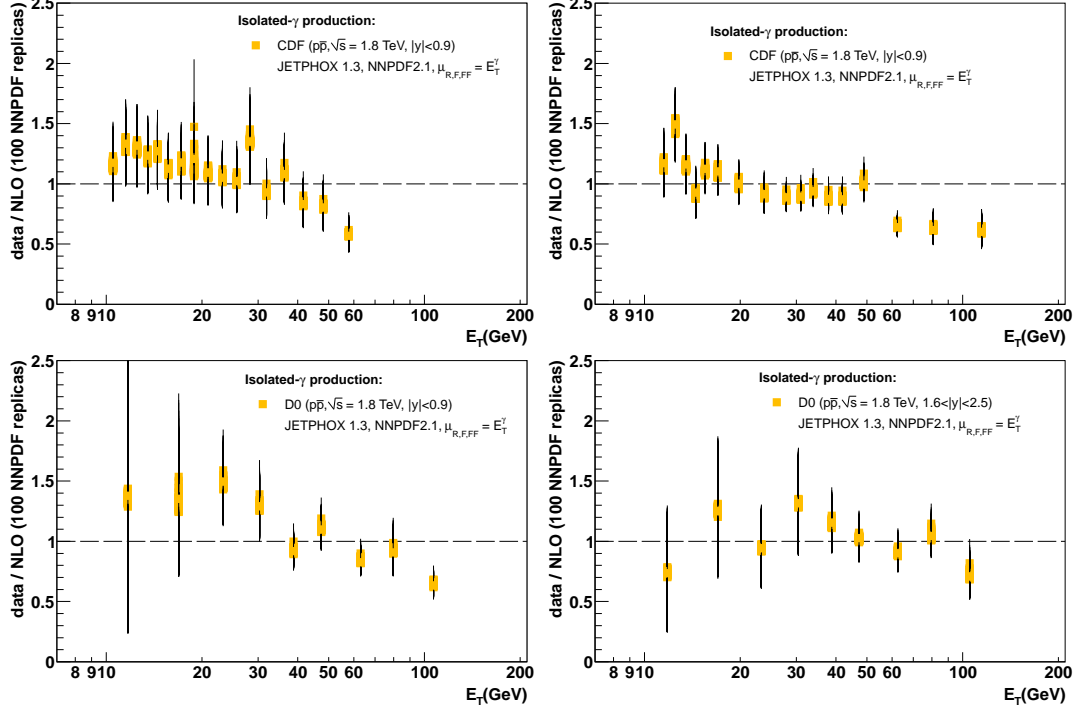


Figure 6: Ratio of isolated-photon data (CDF and D0) and NLO pQCD predictions for $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV at various photon rapidities. The (yellow) band indicates the range of predictions for each one of the 100 NNPDF2.1 replicas, and the bars show the total experimental uncertainty.

the spectral shape seems somewhat different at small photon E_T where the data rise steeper as compared to the theory. The origin of this different shape is still not well understood, although Ref. [24] argues that the discrepancy decreases with a suitable scale choice.

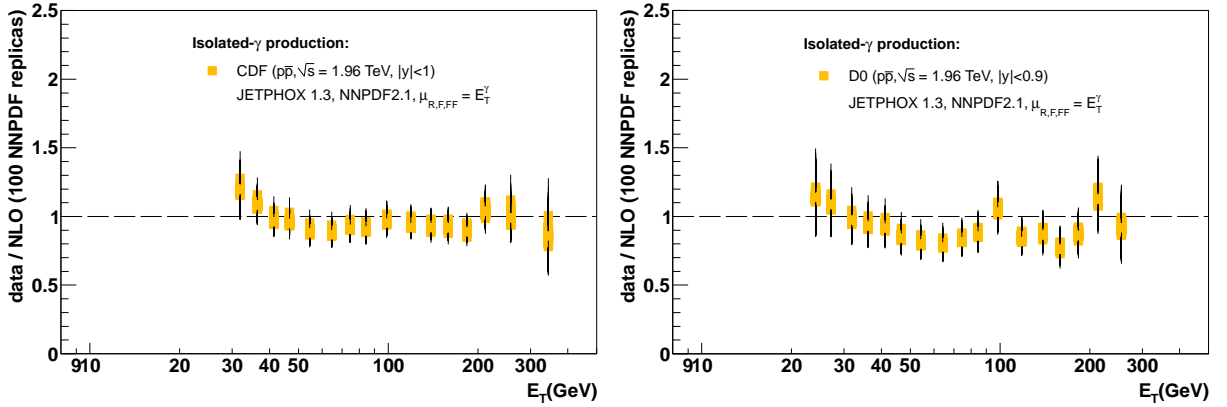


Figure 7: Ratio of isolated-photon data (CDF and D0) and NLO pQCD predictions for $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV at central rapidities. The (yellow) band indicates the range of predictions for each one of the 100 NNPDF2.1 replicas, and the bars show the total experimental uncertainty.

Data-theory comparisons for ATLAS and CMS measurements are shown in Fig. 8 and 9 respectively. We observe a very good agreement for all rapidity ranges, except maybe in a few of the lowest- E_T^γ bins where the central value of the data points tends to undershoot a bit the theoretical

predictions.

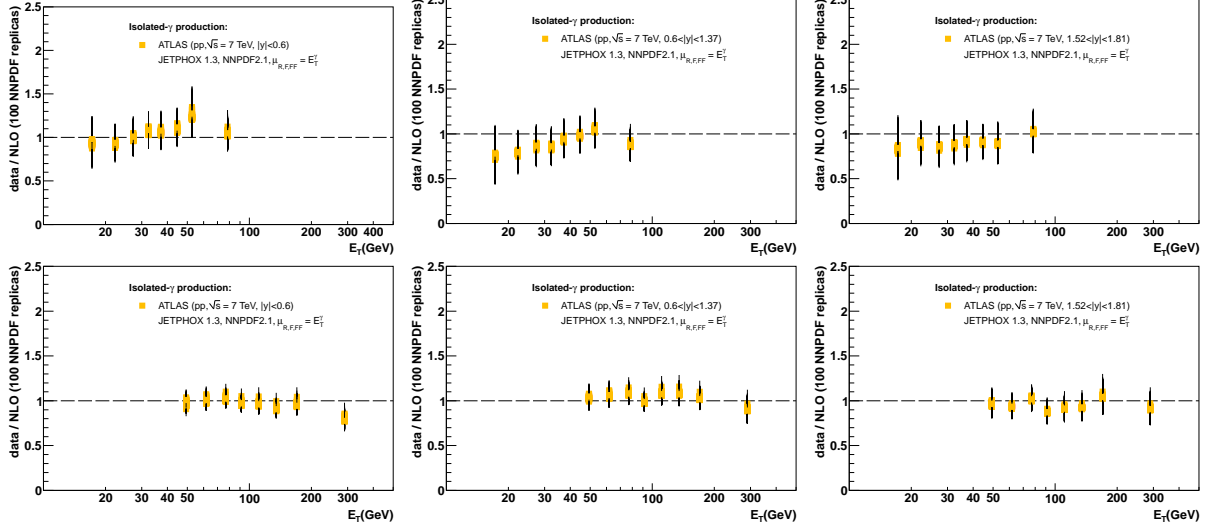


Figure 8: Ratio of ATLAS isolated-photon datasets (880 nb^{-1} , top, and 36 pb^{-1} , bottom) and NLO pQCD predictions in p - p collisions at $\sqrt{s} = 7 \text{ TeV}$ at various photon rapidities. The (yellow) band indicates the range of predictions for each one of the 100 NNPDF2.1 replicas, and the bars show the total experimental uncertainty.

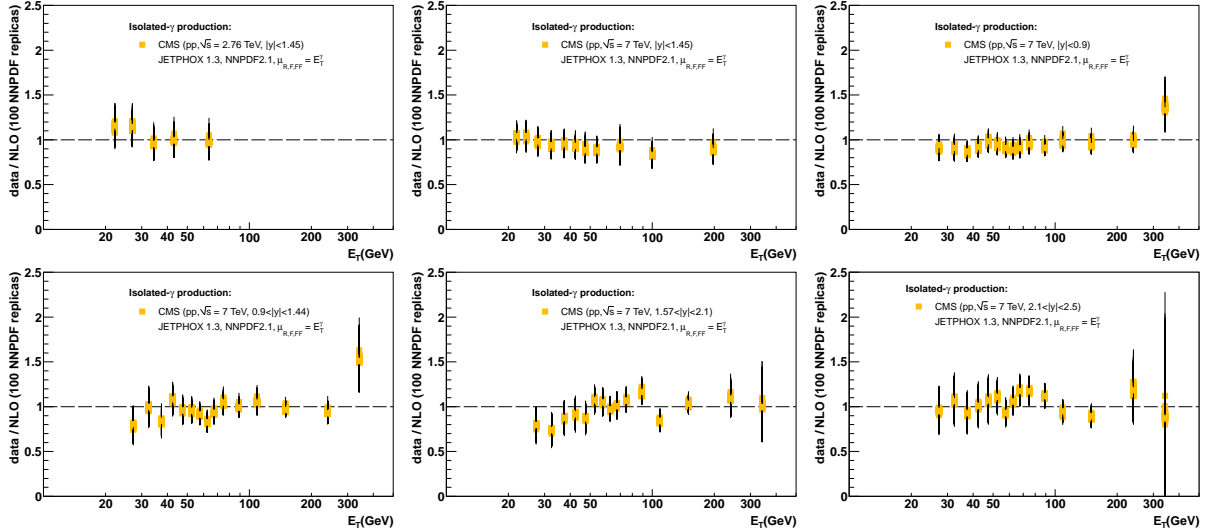


Figure 9: Ratio of CMS isolated-photon data and NLO pQCD predictions in p - p collisions at $\sqrt{s} = 2.76 \text{ TeV}$ (top left) and $\sqrt{s} = 7 \text{ TeV}$ (2.9 nb^{-1} , top-middle, and 36 pb^{-1} , top-right and bottom, datasets) at various photon rapidities. The (yellow) band indicates the range of predictions for each one of the 100 NNPDF2.1 replicas, and the bars show the total experimental uncertainty.

Figure 10 shows the overall comparison of the full collider photon dataset considered in this analysis with the NLO predictions obtained using the central NNPDF2.1 replicas, as a function of x_T . Most of the data/theory ratios are around unity indicating an overall good agreement between NLO pQCD and the experimental measurements. No systematic deviation is observed

in a wide range of x_T corresponding to a wide kinematical coverage of photon E_T , rapidity and collision energy. This is to be contrasted with similar systematics studies [24] that indicate clear data-NLO deviations for the inclusive- γ production in the E706 results at fixed-target energies.

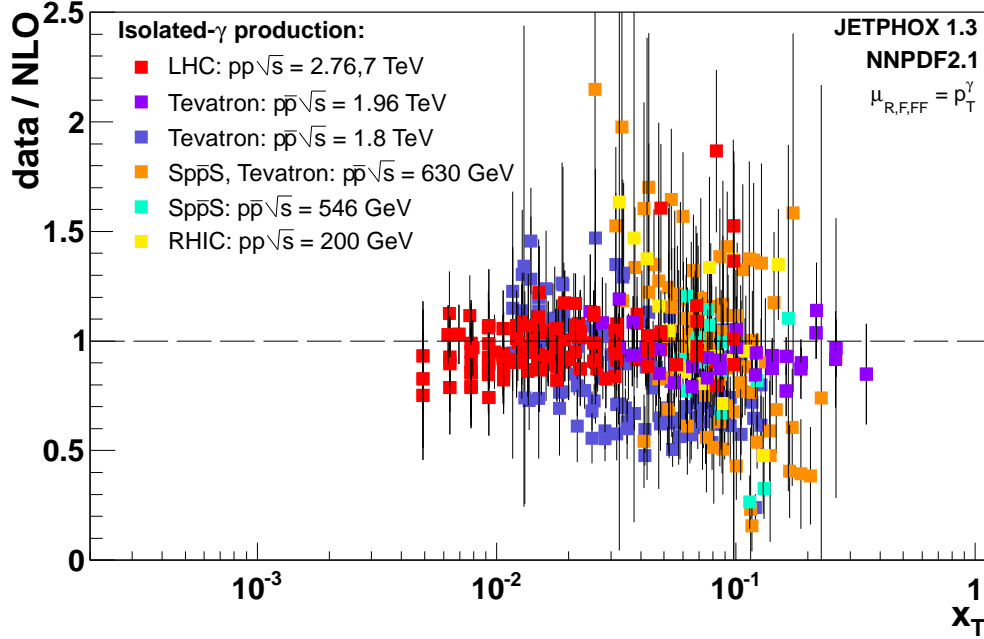


Figure 10: Summary plot for the data/theory ratios for all collider isolated-photon data considered in this analysis, as a function of $x_T = 2E_T^\gamma/\sqrt{s}$. For each system, the NLO prediction used is the one obtained with the central PDF replica of the NNPDF2.1 set. The error bars indicate the total experimental uncertainty.

In Table 1 we quote the average χ^2 over all replicas between each one of the datasets and the NLO calculations. As one can see, for a large majority of cases the agreement is quite good, while in a few cases the χ^2 obtained is rather poor ($\chi^2 \gg 1$). The total initial χ^2 of all the systems considered is $\chi^2 = 1.3$, while after reweighting it decreases to $\chi_{\text{rw}}^2 = 1.1$ for the whole dataset. This global result confirms at the quantitative level that there is a good agreement between NLO pQCD and the experimental results measured at all collider energies.

Looking in more detail, it is worth noticing that for those few systems which are not well reproduced by NLO pQCD, there are always other measurements¹ covering similar $(\sqrt{s}, E_T^\gamma, y_\gamma)$ domains with a good χ^2 . This fact indicates that in such cases the problem is not likely related to the theoretical prediction but of experimental origin. In other words, the issue is not data-theory compatibility but rather an inconsistency problem between measurements covering the same kinematics. Note that in the cases where the χ^2 is poor, the reweighted χ_{rw}^2 is only slightly better, which confirms that the data are not consistent with the theory even after refitting. To quantify better this effect, we can make use of the probability distribution for the rescaling variable α discussed in Sect. 3.2. The last column of Table 2 lists the mean value of the $\mathcal{P}(\alpha)$ distribution², Eq. (5), for all systems. In Fig. 11 we show the $\mathcal{P}(\alpha)$ distribution for those systems

¹The only exception to this are two UA1 measurements at 546 GeV (Fig. 4) for which the large χ^2 is just driven by a single outlier data-point at the highest E_T^γ measured.

²It is easy to check that if the underlying distribution is a χ^2 distribution, the mean value of $\mathcal{P}(\alpha)$ is given by

System	Collab./Experiment	\sqrt{s}	$ y $	E_T^γ range	χ^2	χ_{rw}^2	$\langle\alpha\rangle$
p - p	ATLAS (LHC)	7.	<0.6	15–100	0.2	0.5	0.6
p - p	ATLAS (LHC)	7.	0.6–1.37	15–100	0.5	0.3	0.8
p - p	ATLAS (LHC)	7.	1.52–1.81	15–100	0.3	0.6	0.6
p - p	ATLAS (LHC)	7.	<0.6	45–400	0.7	0.8	1.0
p - p	ATLAS (LHC)	7.	0.6–1.37	45–400	0.3	0.4	0.7
p - p	ATLAS (LHC)	7.	1.52–1.81	45–400	0.3	0.3	0.7
p - p	ATLAS (LHC)	7.	1.81–2.37	45–400	1.6	1.7	1.6
p - p	CMS (LHC)	7.	<1.45	21–300	0.6	0.6	0.8
p - p	CMS (LHC)	7.	<0.9	25–400	1.3	1.1	1.2
p - p	CMS (LHC)	7.	0.9–1.44	25–400	0.8	0.8	1.0
p - p	CMS (LHC)	7.	1.57–2.1	25–400	0.7	0.8	1.0
p - p	CMS (LHC)	7.	2.1–2.5	25–400	0.5	0.5	0.8
p - p	CMS (LHC)	2.76	<1.45	20–80	0.2	0.2	0.7
p - \bar{p}	CDF (Tevatron)	1.96	<1.0	30–400	0.7	0.8	1.3
p - \bar{p}	D0 (Tevatron)	1.96	<0.9	23–300	1.4	1.3	0.9
p - \bar{p}	CDF (Tevatron)	1.8	<0.9	11–132	3.6	2.9	2.1
p - \bar{p}	CDF (Tevatron)	1.8	<0.9	10–65	1.0	1.1	1.1
p - \bar{p}	D0 (Tevatron)	1.8	<0.9	10–140	3.1	2.9	2.1
p - \bar{p}	D0 (Tevatron)	1.8	1.6–2.5	10–140	1.5	0.6	1.9
p - \bar{p}	CDF (Tevatron)	0.63	<0.9	8–38	1.1	1.1	1.3
p - \bar{p}	D0 (Tevatron)	0.63	<0.9	7–50	0.9	1.1	1.3
p - \bar{p}	D0 (Tevatron)	0.63	1.6–2.5	7–50	0.8	0.9	1.2
p - \bar{p}	UA1 (Sp \bar{p} S)	0.63	<0.8	16–100	1.5	1.5	1.3
p - \bar{p}	UA1 (Sp \bar{p} S)	0.63	0.8–1.4	16–70	1.7	1.7	1.5
p - \bar{p}	UA1 (Sp \bar{p} S)	0.63	1.6–3.0	16–70	5.9	4.2	2.7
p - \bar{p}	UA2 (Sp \bar{p} S)	0.63	<0.76	14–92	1.1	1.1	1.2
p - \bar{p}	UA2 (Sp \bar{p} S)	0.63	<0.76	12–83	3.0	2.7	1.9
p - \bar{p}	UA2 (Sp \bar{p} S)	0.63	1.0–1.8	12–51	1.7	1.7	1.5
p - \bar{p}	UA1 (Sp \bar{p} S)	0.546	<0.8	16–51	0.4	0.4	0.8
p - \bar{p}	UA1 (Sp \bar{p} S)	0.546	0.8–1.4	16–46	6.1	5.7	3.3
p - \bar{p}	UA1 (Sp \bar{p} S)	0.546	1.6–3.0	16–38	7.4	6.1	3.6
p - p	PHENIX (RHIC)	0.2	<0.35	3–16	0.6	0.6	0.8

Table 2: Summary of the χ^2 -analysis between NLO pQCD and the world isolated- γ data. For each system we list the initial data-theory χ^2 (6th column), the χ_{rw}^2 (7th column) obtained after including each corresponding dataset via PDF reweighting, and $\langle\alpha\rangle$ (last column), the mean of the associated $\mathcal{P}(\alpha)$ distribution.

in Table 2 with initial $\chi^2 \geq 3$. In all cases $\mathcal{P}(\alpha)$ peaks at a value of α above ~ 2 . This fact points to an underestimation of the experimental uncertainties in the corresponding measurements.

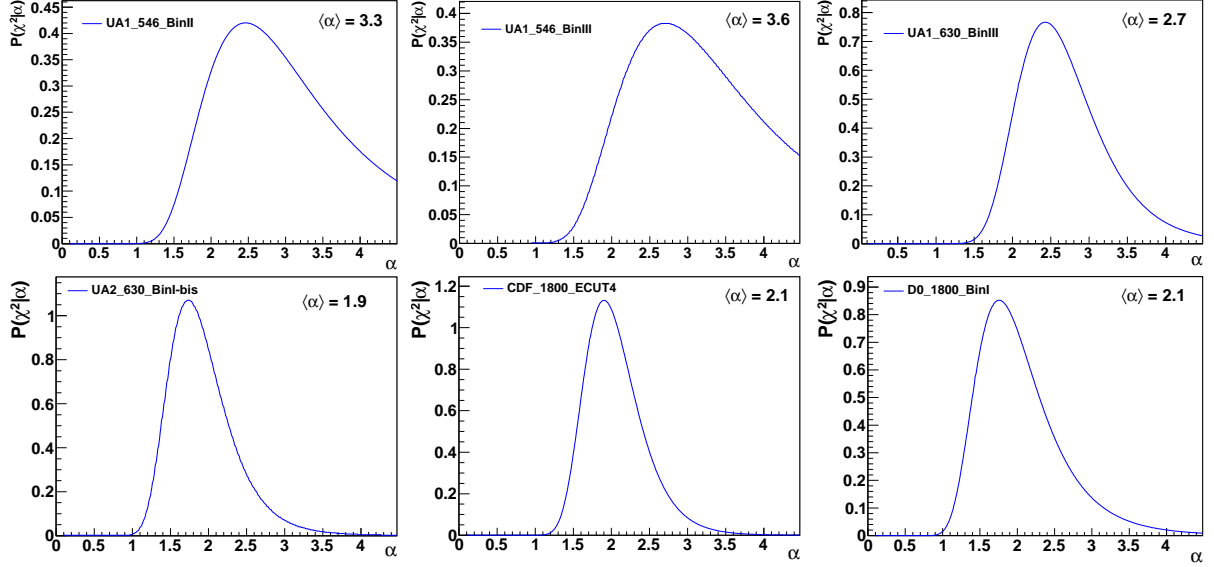


Figure 11: Distribution of the α rescaling variable, Eq. (5), for those datasets that show poor agreement with NLO pQCD. From top-to-bottom and left-to-right, we show two UA1-546 GeV datasets ($\chi^2 = 6.1$ and 7.4 respectively), one UA1 ($\chi^2 = 5.9$) and UA2 ($\chi^2 = 3.0$) datasets at 630 GeV, and finally two Tevatron Run-I CDF and D0 sets ($\chi^2 = 3.6$ and 3.1 respectively).

As a comparison, in Fig. 12 we show the $\mathcal{P}(\alpha)$ distributions for the rapidity bins of the 36 pb^{-1} results from ATLAS and CMS. In all cases $\mathcal{P}(\alpha)$ peaks close to one, confirming the consistency of these datasets with NLO pQCD and the proper estimation of their associated experimental errors.

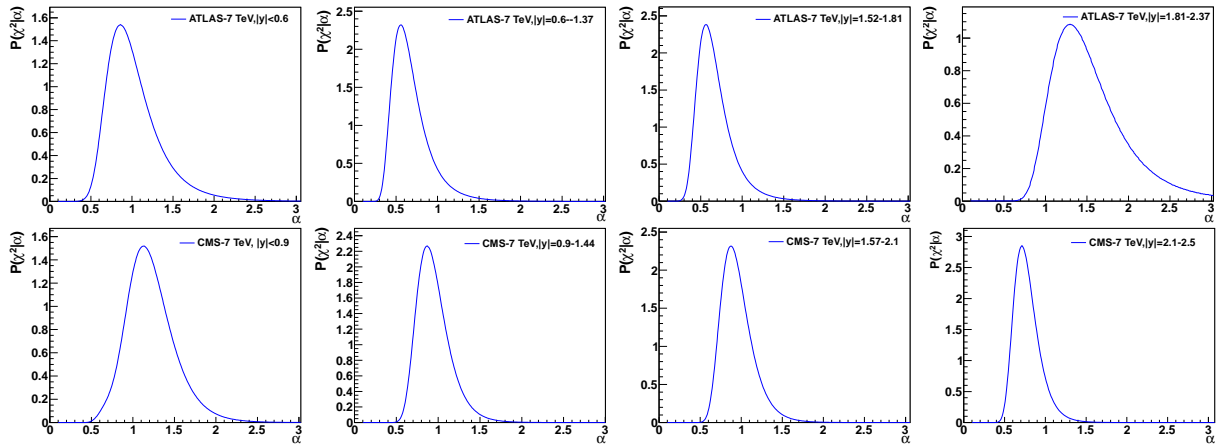


Figure 12: Distribution of the α rescaling variable, Eq. (5), for each one of the rapidity bins of the 36 pb^{-1} datasets from ATLAS (top plots) and CMS (bottom plots).

$\langle \alpha \rangle = 1 + 1/2N_{\text{dat}}$ with N_{dat} the number of data points of that particular system.

4.2 Sensitivity of isolated- γ to the proton PDFs

The dependence of isolated- γ production at different kinematics ranges on the individual flavour of the underlying parton densities can be quantified by computing the correlation coefficient between each one of the light-quark and gluon distributions and the NLO cross sections as discussed in [33]. Correlation profiles are shown in Fig. 13 for four representative values of the photon energy and rapidity typical of the LHC measurements. Isolated photons at central rapidities (top panels) have a dominant sensitivity to $g(x, Q^2)$ in a range between $x = 0.01$ and $x = 0.2$ for increasing photon transverse energy. At forward rapidities (bottom panels) isolated- γ are sensitive to the gluon densities at x values around 10^{-3} for low E_T^γ , but the PDF sensitivity shifts to the valence u -quarks as the photon energy increases. Being able to carry out measurements at even more forward rapidities, e.g. down to $y_\gamma \approx 5$ accessible in LHCb [74] and/or at low- E_T^γ accessible via photon-conversions in ALICE [75], would be very interesting in order to probe x values below 10^{-4} where $g(x, Q^2)$ is only weakly constrained [76] by HERA F_2 , F_L , $F_{2,c}$ structure functions and by the momentum sum-rule.

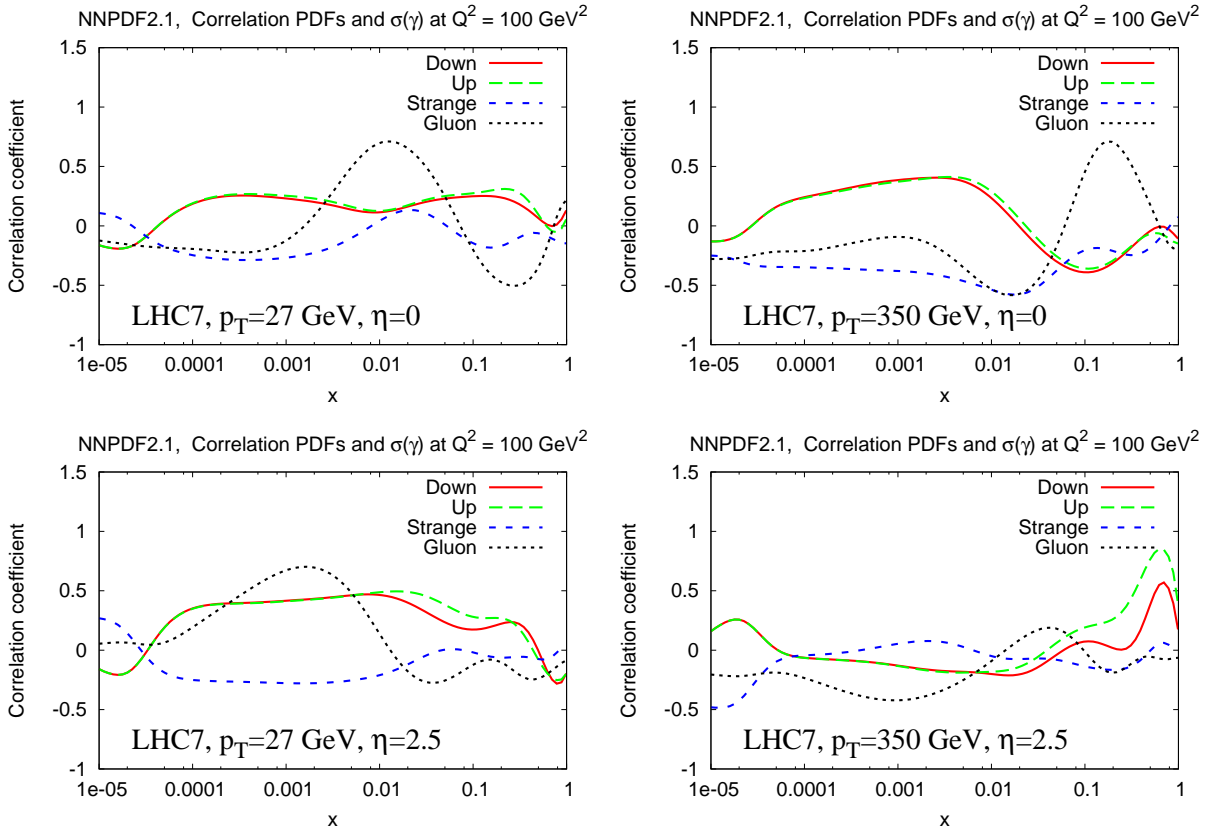


Figure 13: Correlation between the isolated- γ cross section and various flavours of the NNPDF2.1 parton densities for different (E_T^γ, y_γ) kinematical ranges at the LHC. Top (bottom) plots show the correlation at central (forward) rapidities. Left (right) plots show the correlation for typical low (high) E_T^γ values.

In Fig. 14 we plot the sensitivity of the isolated- γ data to the proton gluon at a fixed value of the photon transverse energy ($E_T^\gamma = 27$ GeV) for various collider energies at central (left) and forward (right) rapidities. At midrapidity, the peak of the correlations shifts towards larger x values for decreasing \sqrt{s} , up to about $x \approx 0.1$ at $\mathcal{O}(500$ GeV). At forward rapidities the LHC

data are more sensitive to the gluon density than measurements at Sp̄pS and Tevatron because the former involve p - p collisions (dominated by qg -Compton processes) rather than p - \bar{p} collisions (with $q\bar{q}$ -annihilation playing a larger role, Fig. 1), and because the (E_T^γ, y_γ) phase-space covered by ATLAS and CMS is much larger than that at lower c.m. energies.

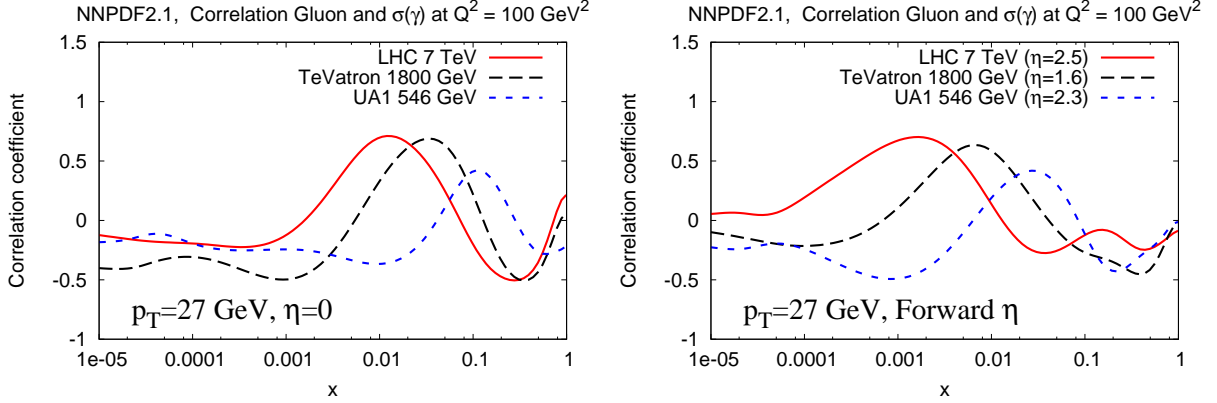


Figure 14: Correlations between the gluon PDF and the isolated- γ cross section at a fixed $E_T^\gamma = 27$ GeV in p - p and p - \bar{p} collisions at central (left) and forward (right) rapidities for different collider energies.

4.3 Impact of isolated-photon data on the gluon

In order to quantify the constraints that the various photon datasets impose on the gluon PDF, the results of the χ^2 -analyses listed in Table 2 are combined into seven groups according to each collider energy : 200, 546, 630 GeV and 1.8, 1.96, 2.76 and 7 TeV. In the grouping, we discard the six datasets with $\chi^2 \geq 3$ which have underestimated experimental uncertainties (Fig. 11) and only result in a loss of accuracy of the reweighting method without any impact on the PDFs whatsoever.

The effective number of replicas N_{eff} after reweighting, Eq. (4), for each collider energy is summarised in Table 3. As expected, the more constraining datasets, i.e. those with smallest N_{eff} , are those from the LHC at 7 TeV. The measurement at 2.76 TeV has only 5 data-points fully consistent with NLO with basically no impact on the PDFs. Of the rest, only the Tevatron Run-II and the 630 GeV datasets seem to have also some, albeit small, reduction of N_{eff} .

\sqrt{s} (TeV)	0.2	0.546	0.630	1.8	1.96	2.76	7
N_{eff}	99.6	99	95	99.8	96	96	87

Table 3: Effective number of replicas N_{eff} , Eq. (4), left for each one of the collider energies considered, after inclusion of the isolated- γ data. The starting number of NNPDF replicas for this analysis is $N_{\text{rep}} = 100$.

In Fig. 15 we show the 100 replica weights, Eq. (3), for three representative collider energies: PHENIX at $\sqrt{s} = 200$ GeV, UA1/UA2/Tevatron at $\sqrt{s} = 630$ GeV, and the LHC data at $\sqrt{s} = 7$ TeV. For the least constraining dataset (the PHENIX data with N_{eff} closer to N_{rep}) all replicas have essentially the same weight. For the other two datasets the distributions become broader, specially for LHC at 7 TeV, showing that the impact of these measurements on the PDFs is larger – some replicas are preferred (larger weight) than others (smaller weight) when

confronted to the photon results.

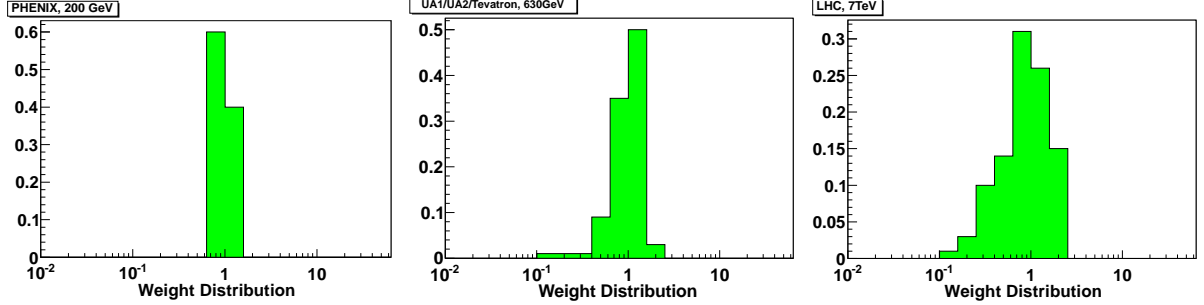


Figure 15: Distribution of NNPDF replicas weights, Eq. (3), for three representative collider energies: PHENIX data at $\sqrt{s} = 200$ GeV (left), $\sqrt{s} = 630$ GeV data (center), and LHC results at $\sqrt{s} = 7$ TeV (right).

The direct quantification of the impact on the $g(x, Q^2)$ distribution is shown in Figs. 16–18 where the NNPDF2.1 NLO gluon is shown before and after including the isolated-photon data with reweighting. In Figs. 16–17 the PDFs are evaluated at a typical LHC scale of $Q = 100$ GeV. We show both the relative improvement (left panels, ratio over the original gluon PDF) as well as the reduction in the absolute PDF errors (right panels). Only the LHC data lead to a significant PDF uncertainty reduction, of up to 20 percent, localised at medium $x \approx 0.002$ to 0.05 (Fig. 18). The Tevatron Run-II and the 630-GeV measurements bring rather small improvements around $x \approx 0.01$ – 0.02 , while the other datasets have negligible impact on the gluon central value and associated uncertainties. This is consistent with the expectation that the effects should be maximal in phase-space regions where the photon production cross section depends more strongly on the gluon distribution (Fig. 14). In the cases where the PDF uncertainties are reduced, the central value of $g(x, Q^2)$ is essentially unaffected (Figs. 16–17 left) indicating that the large- x gluon determined from the Tevatron jet data is consistent with the large- x gluon constrained by the LHC photon results. This is an important cross-check of the validity of pQCD factorisation and of the PDF universality using cross sections measured at hadronic colliders. The effect of the LHC data on the gluon is similar at lower scales, where the impact of photon data is shifted to somewhat larger values of x as dictated by DGLAP evolution, as exemplified in Fig. 18.

The relative reduction of the uncertainties in the gluon distribution is better illustrated in Fig. 19. For three scales $Q = 3.16, 10$, and 100 GeV, the maximum $g(x, Q^2)$ uncertainty reduction is 10–20% in the region $x \approx 0.01$ – 0.05 . For the lower scales, in some small regions of x , the PDF errors can be slightly increased when the photon data are included. This is just a fluctuation effect, due to the limited statistics of the reweighting procedure, that is quickly washed out by DGLAP evolution at higher scales.

As a last check, we have confirmed that the constraints on the quark PDFs from the current isolated- γ data are essentially negligible. This is not unexpected since light-quarks distributions are known more accurately than the gluon PDF in the kinematical region relevant for photon production (Fig. 13). However, the impact of the isolated-photon measurements would certainly be more important in the so-called “collider only” fits [70], where the use of data from colliders alone leads to larger PDF uncertainties on the quark sector (constrained mostly by fixed-target data in pre-LHC PDF sets), since in this case the q, \bar{q} densities could be also directly constrained.

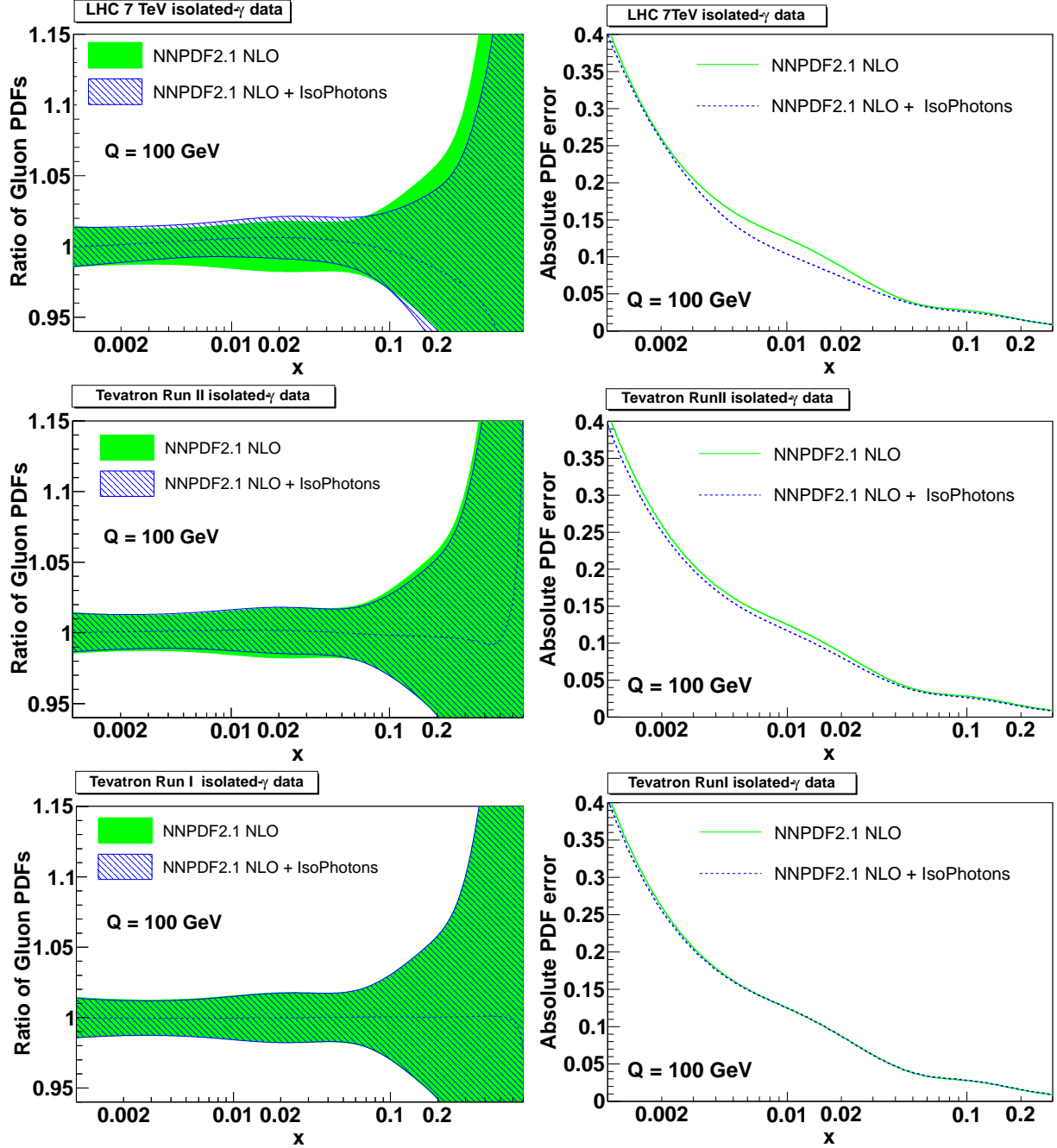


Figure 16: Comparison between the NNPDF2.1 NLO gluon before (green solid band) and after (dashed blue area) inclusion of the isolated- γ data from (top to bottom): LHC-7 TeV, Tevatron Run-II at 1.96 TeV, and Run-II at 1.8 TeV. The left plots show the ratio between the original and the new $g(x, Q^2)$ while the right panels indicates the reduction of absolute $g(x, Q^2)$ uncertainties thanks to the photon data. PDFs are valued at $Q = 100$ GeV, a typical LHC scale.

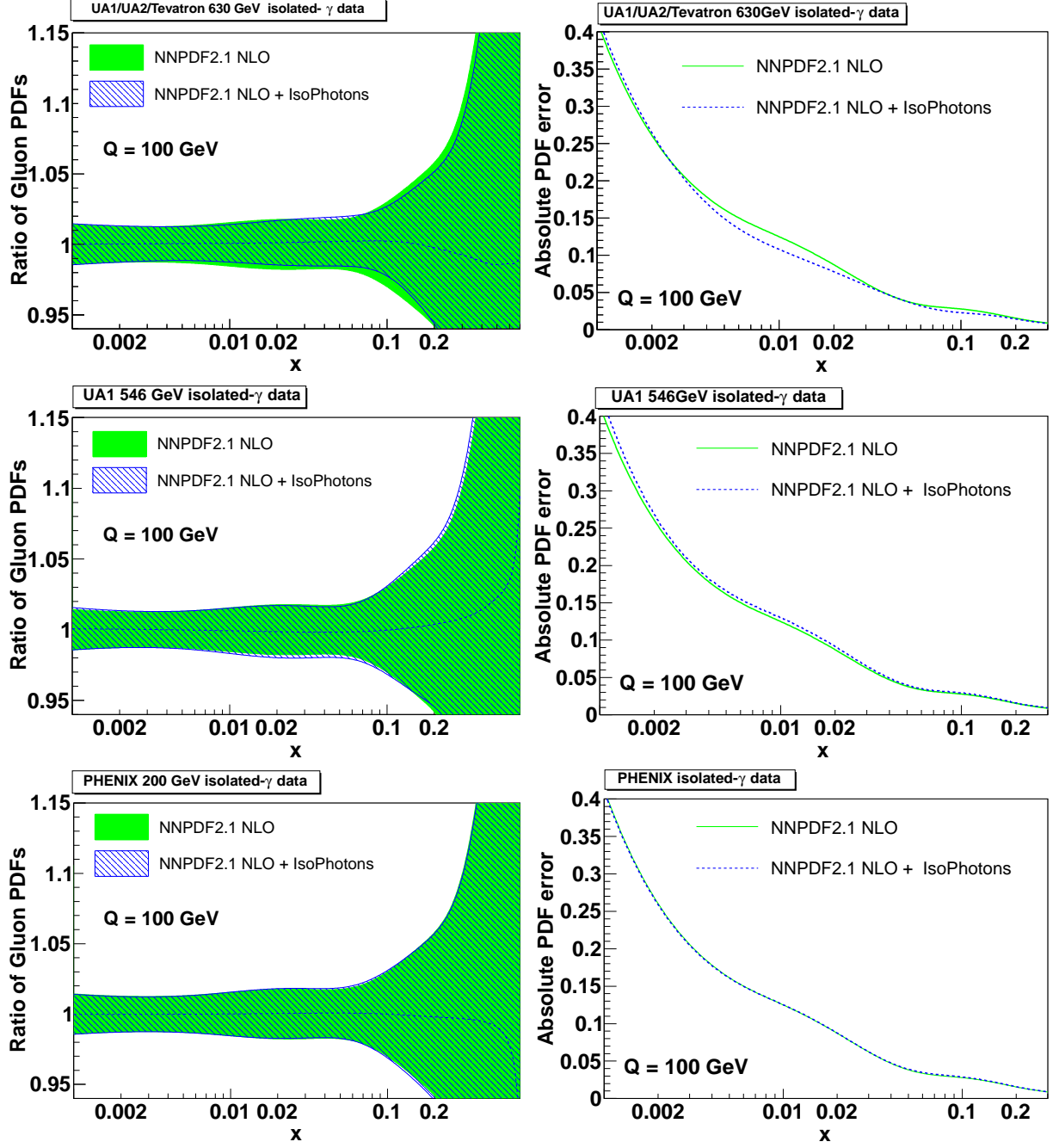


Figure 17: Same as Fig. 16 for the remaining collider data (top to bottom): UA1/UA2/Tevatron at 630 GeV, UA1-546 GeV, and PHENIX-200 GeV.

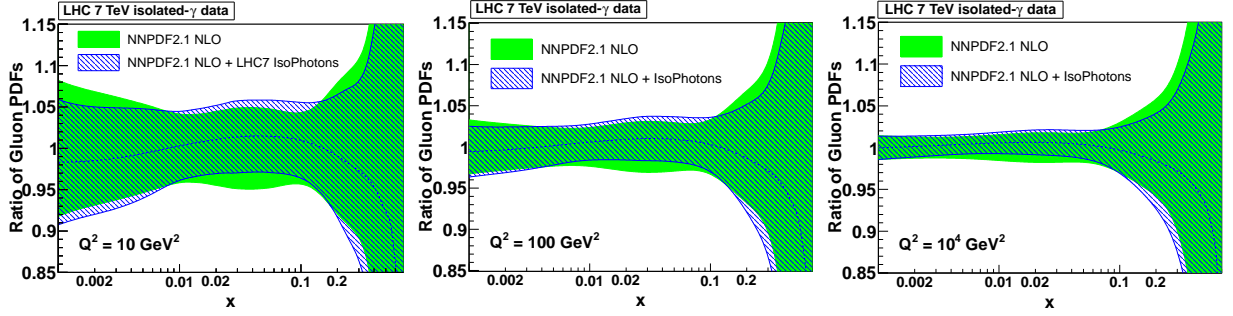


Figure 18: Comparison between the NNPDF2.1 NLO gluon before (green solid band) and after (dashed blue area) inclusion of the LHC-7 TeV isolated- γ data, with PDFs evaluated (left to right) at $Q = 3.16$, 10, and 100 GeV.

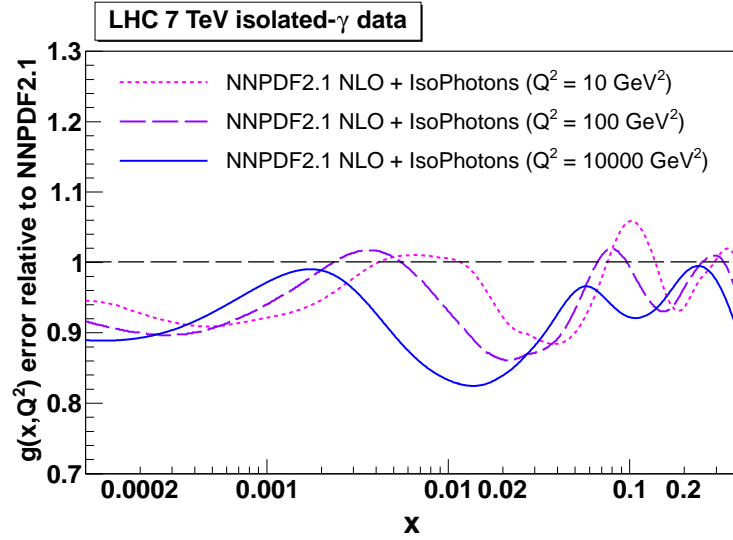


Figure 19: Relative reduction of the NNPDF2.1 NLO gluon distribution uncertainty at scales $Q = 3.16$, 10, and 100 GeV, after inclusion of the LHC isolated-photon data via reweighting.

5 Predictions of the gluon-fusion Higgs boson production cross section

The dominant production channel for the SM Higgs boson at the LHC is gluon fusion [5, 6] and thus theoretical predictions depend strongly on the gluon PDF choice as well as on the associated value of α_s . Recently there has been an intense discussion in the literature [77–79] about the theoretical uncertainties to assign to the $\sigma(gg \rightarrow H)$ cross section at hadron colliders. This has been motivated by the fact that predictions from non-global PDF sets lead to very different results than global PDF fits which are in reasonable agreement among each other [2]. This has of course implications for the current Tevatron and LHC Higgs exclusion limits [80–82]. Dedicated studies [83–85] have shown that, within a global fit, the gluon PDF and the strong coupling are stabilised by the inclusive jet data. It is thus of outmost importance to find additional observables sensitive to $g(x, Q^2)$ to further improve the predictions for gluon-gluon Higgs cross sections, and in this context isolated-photons appear as a promising candidate.

As shown in the previous section, the LHC isolated-photon data leads to a reduction of the $g(x, Q^2)$ uncertainties at around $x \approx 0.02$. Since this kinematical region is relevant for many important gluon-driven processes at the LHC, we study how the PDF uncertainties involved in the determination of the cross sections for various processes, from Higgs boson to top-pair production [86], can be potentially reduced thanks to the inclusion of isolated- γ data. We have decided to consider only 7-TeV photon results since the constraints from measurements at the other c.m. energies are much milder and, in addition, LHC is the only collider for which no dataset needs to be discarded. We show also that the central cross section predictions for Higgs production, mostly driven by the inclusive jet data in global PDF fits, remain stable when isolated-photon data are included.

Process / Cross section	$gg \rightarrow H(120)$	$gg \rightarrow H(160)$	$gg \rightarrow H(200)$	$gg \rightarrow H(500)$
NNPDF2.1	11640 ± 181 fb	6052 ± 103 fb	3494 ± 66 fb	219.3 ± 8.3 fb
NNPDF2.1 + LHC IsoPhotons	11701 ± 140 fb	6073 ± 86 fb	3504 ± 56 fb	218.4 ± 7.6 fb

Process / Cross section	$t\bar{t}$	$t\bar{t}H(120)$	$WH(120)$	$ZH(120)$
NNPDF2.1	162 ± 51 pb	114 ± 5 fb	447 ± 9 fb	364 ± 6 fb
NNPDF2.1 + LHC IsoPhotons	162 ± 47 pb	113 ± 4 fb	448 ± 9 fb	365 ± 6 fb

Table 4: NLO cross sections in p - p collisions at 7 TeV obtained with MCFM using the NNPDF2.1 set, before and after including the LHC isolated-photon data. Top: Higgs production in gluon-gluon fusion for different values of M_H . Bottom: QCD top-pair production, and Higgs production ($M_H = 120$ GeV) associated with top-pairs and electroweak bosons.

The production cross sections for the following processes have been computed at NLO³ with the MCFM code [87–89] for p - p collisions at $\sqrt{s} = 7$ TeV: (i) Higgs boson in gluon fusion ($gg \rightarrow H$) for masses $M_H = 120, 160, 200, 300$ and 500 GeV, Higgs boson in association with (ii) top quark pairs ($t\bar{t}H$), and (iii) W and Z (WH , ZH); and (iv) top quark pair production ($t\bar{t}$). All these processes are, either as signal or as background, relevant for Higgs searches at the LHC. The results for the cross sections obtained with the NNPDF2.1 NLO set, before and after including the LHC isolated photon data, are summarised in Table 4. The results for Higgs boson production

³The NNLO/NLO K-factor, when available, is known to depend mildly on the choice of PDF. Therefore, for the study of interest here (impact of the PDF error reduction on LHC cross sections) NLO calculations are sufficient.

in gluon-gluon fusion are also shown in Fig. 20. While the most updated searches exclude the SM Higgs boson with masses outside the 115–127 GeV range [81,82], we nevertheless show a wider mass range to illustrate the impact of the photon data.

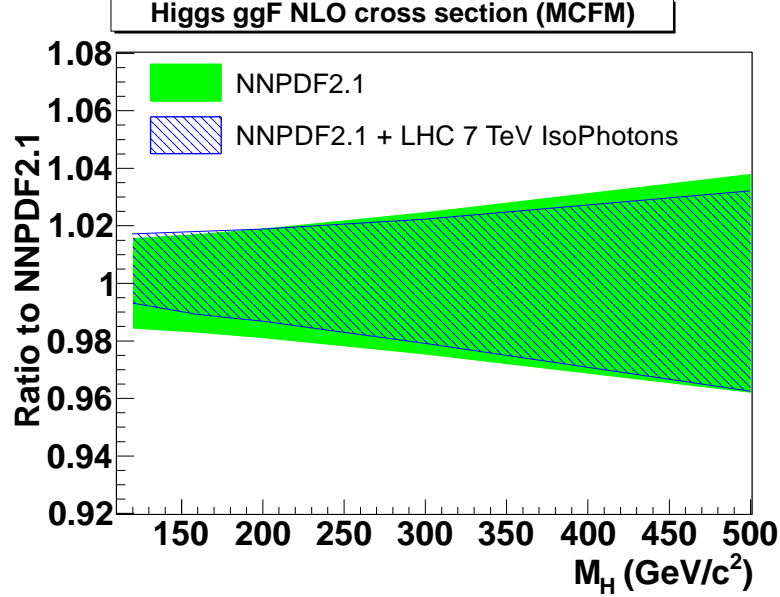


Figure 20: Ratio of Higgs production gluon-fusion cross sections with NNPDF2.1 NLO PDFs before and after including the LHC isolated-photon data.

As one can see from Table 4 and Fig. 20, the inclusion of isolated-photon results in the NNPDF analysis leads to reduced PDF uncertainties in the gluon-gluon Higgs production cross section, which is maximal for low Higgs masses (in the region not yet excluded by the ATLAS and CMS limits) and can be as large as 20%. For other processes the improvements are much more modest. Indeed, processes with a pair of top-quarks are produced with much larger virtualities than a (low-mass) Higgs and have reduced gluon uncertainties. The same holds for processes with associated electroweak boson production which are dominated by quark PDFs. In all cases, the central prediction is in good agreement with the reference NNPDF2.1 results, and thus consistent with the information obtained from the inclusive jet data in a global fit. All in all, these results demonstrate that isolated-photon data can be useful to reduce the theoretical (PDF) uncertainties in the cross section for the many gluon-induced processes at the LHC.

6 Summary and outlook

Using up-to-date NLO pQCD theoretical calculations for isolated-photon production, as implemented in the JETPHOX program combined with the NNPDF2.1 parton densities and its associated PDF reweighting technique, we have quantified the impact on the gluon density of all the existing isolated- γ data measured in p - p and p - \bar{p} collisions at collider energies. The main results of this work can be summarised as follows:

- NLO pQCD provides generally a good description of the isolated-photon measurements from 200 GeV up to 7 TeV in a wide kinematic range of photon transverse energies and rapidities. A few outliers exist which can be identified as arising from problems in experimental datasets with underestimated systematic uncertainties.
- The ATLAS and CMS measurements at the LHC, amounting to more than one hundred data-points, are consistent with each other and with pQCD and precise enough to moderately constrain the gluon PDF in the region $x \approx 0.002 - 0.05$ in a large range of scales $Q^2 = 10-10^4$ GeV². The central value of the gluon distribution is unmodified but its uncertainty is reduced by up to 20% around $x \approx 0.02$.
- New PDFs including the LHC photon data lead to improved predictions for low mass Higgs production in the gluon-fusion channel, with central values of the theoretical cross section unmodified but with associated PDF uncertainties decreased by as much as 20%.

These results confirm that there is no reason not to include isolated-photon data measured at LHC energies into global-fit PDF analyses, on a similar footing as the current inclusive jet production data. The main difficulty to systematically carry out such a program is the slowness of the NLO pQCD codes to compute the corresponding theoretical predictions, and thus an eventual fast interface *à la* APPLGRID or FASTNLO would be very advantageous.

Various possible future developments of this study are outlined next. First, from the experimental side, measurements of the inclusive single isolated-photon spectra at high- E_T^γ exploiting the much larger luminosity collected in 2011 by the ATLAS and CMS collaborations at 7 TeV, as well as data at 8 TeV in 2012 and eventually at 14 TeV would be very beneficial. The addition of a few hundred new data-points over a large kinematic range, possibly including the enhanced E_T^γ - y_γ range covered by ALICE and LHCb, will certainly lead to even stronger constraints on the gluon density. The measurement of ratios of photon cross sections at 8 (or 14) to 7 TeV, where part of the experimental uncertainties cancel out, and/or of the double-differential γ -jet spectra (with constrained kinematics by the concurrent measurement of the photon and jet) can also provide new important constraints on $g(x, Q^2)$. Second, at the lowest end of collider energies, the possibility of measuring isolated- γ production in the full range of the RHIC c.m. energies ($\sqrt{s} \approx 20-500$ GeV) would be also very useful to clarify for once the long-standing disagreement between fixed-target data (as well as the oldest ISR and a few of the Sp \bar{p} S results) and NLO calculations. In all cases, the use of a smooth-cone prescription for the photon isolation [90] which can effectively remove any remaining fragmentation-photon component would be helpful to further eliminate theoretical uncertainties. Third, it would be also important that future photon results provide the full covariance matrix, just as recent LHC measurements of electroweak boson [91] and inclusive jet [92] production have done, as this would improve the quantitative treatment of these datasets in PDF analyses.

On the theoretical side, it would be useful to carry out a more quantitative study of the possible impact of scale variations on the photon cross sections. It is conceivable that part of the impact of the photon data is washed out once higher-order uncertainties are accounted for. This situation is common to all hadronic production data from colliders (notably, inclusive jets) included in global fits. One possibility to reduce this problem would be to include threshold re-summation corrections. In any case, such a scale variation analysis should be consistently carried out together with all the other datasets. Once photon data are included in PDF analyses, it could be also important to accurately and consistently determine the strong coupling in the framework of a PDF fit, just as inclusive jet production is now instrumental in that regard [83, 85, 93] as compared to DIS-only fits. This is so because the gluon is loosely constrained in DIS-only fits, where it has a runaway direction.

As a conclusion, we have shown that the available isolated-photon data provides constraints on the gluon PDFs and thus on many relevant LHC processes, most importantly Higgs production in gluon-gluon fusion. Given that even more precise data as well as theoretical improvements will be available in the next future, we see no objection why isolated-photon data should not become integral part of future global QCD analyses.

Acknowledgments

D. d'E. is grateful to François Arleo and Jean-Philippe Guillet for many discussions and for implementing various improvements in the JETPHOX code which facilitated significantly the computation of many thousands of theoretical spectra. D. d'E. expresses his gratitude to Raphaëlle Ichou for her help on the compilation of the experimental isolated-photon data. Valuable exchanges with Graeme Watt on PDF and NLO calculations are also acknowledged. The research of J. R. has been supported by a Marie Curie Intra-European Fellowship of the European Community's 7th Framework Programme under contract number PIEF-GA-2010-272515.

References

- [1] A. De Roeck and R. Thorne, *Prog. Part. Nucl. Phys.* 66 (2011) 727, 1103.0555.
- [2] G. Watt, *JHEP* 1109 (2011) 069, arXiv:1106.5788.
- [3] G. Watt, (2012), arXiv:1201.1295.
- [4] H1 and ZEUS Collab.s, F.D. Aaron et al., *JHEP* 01 (2010) 109, arXiv:0911.0884.
- [5] LHC Higgs Cross Section Working Group, S. Dittmaier et al., (2011), arXiv:1101.0593.
- [6] LHC Higgs Cross Section Working Group, S. Dittmaier et al., (2012), arXiv:1201.3084.
- [7] J.F. Owens, *Rev. Mod. Phys.* 59 (1987) 465.
- [8] P. Aurenche et al., *Phys. Rev. D* 39 (1989) 3275.
- [9] W. Vogelsang and A. Vogt, *Nucl. Phys. B* 453 (1995) 334, hep-ph/9505404.
- [10] CTEQ Collab., H.L. Lai et al., *Phys. Rev. D* 55 (1997) 1280, hep-ph/9606399.

- [11] A.D. Martin et al., Eur. Phys. J. C14 (2000) 133, hep-ph/9907231.
- [12] UA6 Collab., M. Werlen et al., Phys. Lett. B452 (1999) 201.
- [13] Fermilab E706 Collab., G. Alverson et al., Phys. Rev. D48 (1993) 5.
- [14] Fermilab E706 Collab., L. Apanasevich et al., Phys. Rev. Lett. 81 (1998) 2642, hep-ex/9711017.
- [15] Fermilab E706 Collab., L. Apanasevich et al., Phys. Rev. D70 (2004) 092009, hep-ex/0407011.
- [16] H. Baer, J. Ohnemus and J.F. Owens, Phys. Rev. D42 (1990) 61.
- [17] P. Aurenche et al., Nucl. Phys. B399 (1993) 34.
- [18] L. Gordon and W. Vogelsang, Phys. Rev. D50 (1994) 1901.
- [19] E. Laenen, G. Oderda and G.F. Sterman, Phys. Lett. B438 (1998) 173, hep-ph/9806467.
- [20] S. Catani, M.L. Mangano and P. Nason, JHEP 07 (1998) 024, hep-ph/9806484.
- [21] N. Kidonakis and J.F. Owens, Phys. Rev. D63 (2001) 054019, hep-ph/0007268.
- [22] N. Kidonakis and J.F. Owens, Int. J. Mod. Phys. A19 (2004) 149, hep-ph/0307352.
- [23] D. de Florian and W. Vogelsang, Phys. Rev. D72 (2005) 014014, hep-ph/0506150.
- [24] P. Aurenche et al., Phys. Rev. D73 (2006) 094007, hep-ph/0602133.
- [25] W. Vogelsang and M.R. Whalley, J. Phys. G23 (1997) A1.
- [26] P. Aurenche et al., Eur. Phys. J. C9 (1999) 107, hep-ph/9811382.
- [27] R. Ichou and D. d’Enterria, Phys. Rev. D82 (2010) 014015, arXiv:1005.4529.
- [28] L. Apanasevich et al., Phys. Rev. D59 (1999) 074007, hep-ph/9808467.
- [29] A. Gehrmann-De Ridder and E.W.N. Glover, Eur. Phys. J. C7 (1999) 29, hep-ph/9806316.
- [30] M. Gluck, E. Reya and A. Vogt, Phys. Rev. D48 (1993) 116.
- [31] L. Bourhis, M. Fontannaz and J.P. Guillet, Eur. Phys. J. C2 (1998) 529, hep-ph/9704447.
- [32] P. Aurenche et al., http://lappweb.in2p3.fr/lapth/PHOX_FAMILY/jetphox.html.
- [33] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B849 (2011) 296, arXiv:1101.1300.
- [34] ATLAS Collab., G. Aad et al., Phys. Rev. D83 (2011) 052005, arXiv:1012.4389.
- [35] ATLAS Collab., G. Aad et al., Phys. Lett. B706 (2011) 150, arXiv:1108.0253.
- [36] CMS Collab., S. Chatrchyan et al., Phys. Rev. D84 (2011) 052011, arXiv:1108.2044.
- [37] CMS Collab., V. Khachatryan et al., Phys. Rev. Lett. 106 (2011) 082001, arXiv:1012.0799.
- [38] CMS Collab., S. Chatrchyan et al., (2012), arXiv:1201.3093.
- [39] T. Carli et al., Eur. Phys. J. C66 (2010) 503, arXiv:0911.2985.

- [40] T. Kluge, K. Rabbertz and M. Wobisch, (2006), hep-ph/0609285.
- [41] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B849 (2011) 112, arXiv:1012.0836.
- [42] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B855 (2012) 608, arXiv:1108.1758.
- [43] D. d’Enterria and R. Ichou, PoS (EPS-HEP2011) (2011) 291.
- [44] CDF Collab., T. Aaltonen et al., Phys. Rev. D80 (2009) 111106, arXiv:0910.3623.
- [45] D0 Collab., V.M. Abazov et al., Phys. Lett. B639 (2006) 151, hep-ex/0511054.
- [46] CDF Collab., D. Acosta et al., Phys. Rev. D65 (2002) 112003, hep-ex/0201004.
- [47] CDF Collab., D. Acosta et al., Phys. Rev. D70 (2004) 074008, hep-ex/0404022.
- [48] CDF Collab., F. Abe et al., Phys. Rev. Lett. 73 (1994) 2662.
- [49] D0 Collab., B. Abbott et al., Phys. Rev. Lett. 84 (2000) 2786, hep-ex/9912017.
- [50] D0 Collab., S. Abachi et al., Phys. Rev. Lett. 77 (1996) 5011, hep-ex/9603006.
- [51] D0 Collab., V.M. Abazov et al., Phys. Rev. Lett. 87 (2001) 251805, hep-ex/0106026.
- [52] UA1 Collab., C. Albajar et al., Phys. Lett. B209 (1988) 385.
- [53] UA2 Collab., J. Alitti et al., Phys. Lett. B288 (1992) 386.
- [54] UA2 Collab., R. Ansari et al., Z. Phys. C41 (1988) 395.
- [55] PHENIX Collab., S.S. Adler et al., Phys. Rev. Lett. 98 (2007) 012002, hep-ex/0609031.
- [56] F. Arleo et al., Phys. Rev. Lett. 105 (2010) 062002, arXiv:0911.4604.
- [57] P. Aurenche et al., Phys. Lett. B140 (1984) 87.
- [58] F. Aversa et al., Nucl. Phys. B327 (1989) 105.
- [59] G. Diana, Nucl. Phys. B824 (2010) 154, arXiv:0906.4159.
- [60] G. Diana, J. Rojo and R.D. Ball, Phys. Lett. B693 (2010) 430, arXiv:1006.4250.
- [61] T. Becher and M.D. Schwartz, JHEP 02 (2010) 040, arXiv:0911.0681.
- [62] J.H. Kuhn et al., JHEP 03 (2006) 059, hep-ph/0508253.
- [63] D. Bourilkov, R.C. Group and M.R. Whalley, (2006), hep-ph/0605240.
- [64] G.D. Lafferty and T.R. Wyatt, Nucl. Instrum. Meth. A355 (1995) 541.
- [65] NNPDF Collab., L. Del Debbio et al., JHEP 03 (2005) 080, hep-ph/0501067.
- [66] NNPDF Collab., L. Del Debbio et al., JHEP 03 (2007) 039, hep-ph/0701127.
- [67] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B809 (2009) 1, arXiv:0808.1231.
- [68] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B823 (2009) 195, arXiv:0906.1958.

- [69] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B838 (2010) 136, arXiv:1002.4407.
- [70] NNPDF Collab., R.D. Ball et al., Nucl. Phys. B855 (2012) 153, arXiv:1107.2652.
- [71] G. Watt, Talk at the PDF4LHC workshop, CERN, <https://indico.cern.ch/materialDisplay.py?contribId=3&materialId=slides&confId=145744>, 2011.
- [72] M. Dittmar et al., editors, Parton Distributions (HERA and the LHC Workshop Proceedings, 2009) chap. 3.2: Experimental Error Propagation, arXiv:0901.2504.
- [73] NNPDF Collab., R.D. Ball et al., JHEP 05 (2010) 075, arXiv:0912.2276.
- [74] LHCb Collab., F. De Lorenzi, (2010), arXiv:1011.4260.
- [75] ALICE Collab., A. Marin, Eur. Phys. J. C61 (2009) 735.
- [76] J. Rojo and F. Caola, (2009), arXiv:0906.2079.
- [77] J. Baglio and A. Djouadi, JHEP 03 (2011) 055, arXiv:1012.0530.
- [78] J. Baglio et al., Phys. Lett. B699 (2011) 368, arXiv:1101.1832.
- [79] S. Alekhin, J. Blumlein and S. Moch, (2011), arXiv:1101.5261.
- [80] CDF and D0 Collabs., (2011), arXiv:1107.5518, EPS 2011 Proceeds.
- [81] ATLAS Collab., G. Aad et al., (2012), arXiv:1202.1408.
- [82] CMS Collab., S. Chatrchyan et al., (2012), arXiv:1202.1488.
- [83] R. Thorne and G. Watt, JHEP 1108 (2011) 100, arXiv:1106.5789.
- [84] NNPDF Collab., R.D. Ball et al., Phys. Lett. B704 (2011) 36, arXiv:1102.3182.
- [85] NNPDF Collab., R.D. Ball et al., Phys. Lett. B707 (2012) 66, arXiv:1110.2483.
- [86] F. Demartin et al., Phys. Rev. D82 (2010) 014002, arXiv:1004.0962.
- [87] MCFM, <http://mcfm.fnal.gov>.
- [88] J. Campbell, R.K. Ellis and F. Tramontano, Phys. Rev. D70 (2004) 094012, hep-ph/0408158.
- [89] J. Campbell and R.K. Ellis, Phys. Rev. D65 (2002) 113007, hep-ph/0202176.
- [90] S. Frixione, Phys. Lett. B429 (1998) 369, hep-ph/9801442.
- [91] ATLAS Collab., G. Aad et al., (2011), arXiv:1109.5141.
- [92] ATLAS Collab., G. Aad et al., (2011), arXiv:1112.6297.
- [93] S. Lionetti et al., Phys. Lett. B701 (2011) 346, arXiv:1103.2369.